THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *11/12/2017* **par :**

Bertrand JAYLES

EFFECTS OF INFORMATION QUANTITY AND QUALITY ON COLLECTIVE DECISIONS IN HUMAN GROUPS

JURY

Astrid HOPFENSITZ	Maître de conférence	Examinateur
Franck SCHWEITZER	Professeur	Examinateur
Bertrand GEORGEOT	Directeur de Recherche	Examinateur
Bertrand JOUVE	Directeur de Recherche	Examinateur
Jean-Pierre NADAL	Directeur de Recherche	Examinateur
Pablo JENSEN	Directeur de Recherche	Rapporteur
Alain BARRAT	Directeur de Recherche	Rapporteur
Clément SIRE	Directeur de Recherche	Directeur de Thèse
Guy THERAULAZ	Directeur de Recherche	Invité

École doctorale et spécialité :
SDM : Physique de la matière - CO090

Unité de Recherche :
Laboratoire de Physique Théorique (CNRS UMR 5277) et Centre de Recherche sur la Cognition Animale (CNRS UMR 5169)

Directeur(s) de Thèse :
Clément SIRE et *Guy THERAULAZ*

Rapporteurs :
Alain BARRAT et *Pablo JENSEN*

Acknowledgements

A bit over seven years ago, I started an undergraduate program in Physics with the definite objective to get a PhD degree in fundamental theoretical Physics. However, somewhere along the line, an incredible opportunity occurred to me, that steered me a bit away from my initial objective: after my master degree was complete, I stumbled upon an offer for a three years PhD project on "collective decision-making in human groups". The topic was described as the application of theoretical physics tools to sociological and ethological questions, and I found the subject so fascinating that I jumped at the occasion! I didn't even know such a discipline existed, or I would have searched for it more actively. I am deeply grateful to Clément Sire and Guy Theraulaz, and through them to Paul Sabatier University and the National Center for Scientific Research (CNRS) for accepting my application, thus opening a wide new world to me.

As a theoretical physicist by training, I had no background whatsoever in ethology or social sciences (apart from leisure reading), a near-to-zero level in scientific programming, and never run an experiment in my entire life. I was not even on the statistics "side" of physics. This PhD was therefore an amazing opportunity to learn a large amount of new things in various domains. The inherent diversity of the subject itself made me go through a wide range of literature, from psychology to ethology, from sociology to economics, and more. Being part of the Collective Animal Behavior team at Toulouse University helped me a lot in this respect, and I am thankful to its members Vincent Fourcassié, Richard Bon, Christian Jost, Pierre Moretto, Gérard Latil and Mathieu Moreau, for their suggestions, help and friendship.

I also considerably improved my level in scientific programming, and discovered the pleasures of data analysis, which was totally absent of my curriculum. I benefited from the expertise of colleagues and friends to learn several languages (Python, R, Fortran77, C++, Z-Tree), in particular Lionel Lacombe, Valentin Lecheval and Olivier Gauthé, whom I thank warmly.

While I never thought – again as a theoretical physicist – I would one day have to perform an experiment, I actively participated in five during this PhD, and experienced how difficult it is to think of every detail. The time and efforts spent on designing, building, preparing and running experiments should not be underestimated, and I hope the experimental aspects are emphasized enough in this thesis. At the same time, it was a far nicer and more rewarding experience than I would have ever imagined, and I definitely intend to continue combining theoretical and experimental approaches in my future research. Big thanks to Matthieu Roy, Gilles Tredan, Roberto Pasqua and Christophe Zanon for providing a sophisticated tracking system for our experiments with pedestrians, and handling it throughout the experiments.

Because this PhD was intrinsically multi-disciplinary, I had the chance to be part of several collaborations. In particular, my supervisors and I were involved in a project called MuSE (Multi-disciplinary Study of Emergent phenomena) which aims to bring together expertise from biology, economics, physics, mathematics and computer science to study self-organisation in social and economic systems. The MuSE project funded all our experiments, for which I am deeply grateful. Thanks in particular to Adrien Blanchet, Paul Seabright, Astrid Hopfensitz

and Frédéric Amblard for valuable discussions and comments. I also thank MuSE for partly funding two others collaborations I had with Tokyo and New York Universities, which will leave indelible personal and professional memories:

- I spent July and August 2015 in Pr. Tatsuya Kameda's laboratory in Tokyo University, where I learned the language Z-Tree, and how to design and run an experiment. Many thanks to Pr. Kameda for his kind invitation, to his at the time PhD student Hye-rin Kim who was the actual leader of the first experiment I took part in, graduate student Yoshimatsu Saito and undergraduate student Atsushi Ueshima for their kindness and general support in what will remain one of the most extraordinary experience of my life, at every level.
- The following year, I spent June and July at New-York University for an internship with Pr. Bruno Gonçalves, which was also undoubtedly one of the most exciting and rewarding experience of my life. I learned the Python language, and developed data mining and complex networks skills, which I believe are very important in my research area, and more specifically for future research I intend to carry out. Deep thanks to Bruno for offering me this incredible opportunity and his help in learning new skills.

I express my deep gratitude to the NEXT organization for funding the most part of these trips, as well as the COMUE and the ATUPS who covered the rest.

I want to give special thanks to several people who have helped me a lot throughout the administrative issues I had to face. The first ones go to Malika Bentour, our beloved secretary at LPT (Laboratoire de Physique Théorique), who made my life much easier as a PhD student. Thanks also to Patricia Bordais, to Corine Delesalle and Floriane Abo from NEXT, and to Caroline Mondoulet and Maryline Fautrier from MuSE, who have all been very kind to me and patient in guiding me throughout all the missions paperwork.

I will not remember this PhD as being only a rewarding work experience, but also as an amazing opportunity to meet a lot of people from very different horizons – especially since I was part of two different labs – some of whom have become close friends. I want to thank them for making these three years in Toulouse a delightful time, in particular Valentin, Rémi, Guillaume, Lionel, Benjamin, Julie, Lijie, Li, Hugo, Kevin, Hélène, Natasha, Maxime, Karim, Nicolas, Mohammad, and many others.

Special acknowledgements for my colleague and friend Ramón Escobedo, post-doc at CRCA (Centre de Recherche sur la Cognition Animale), who participated closely in all projects, and provided a considerable help. Thanks also to the jury members, and the referees in particular, for their useful comments and insights that helped me make this thesis better.

At the time of writing these lines, I have been hired as a post-doctoral fellow at the Center for Adaptive Rationality (ARC) at Max Planck Institute for Human Development in Berlin, Germany. I want to thank my supervisor Ralf Kurvers for allowing me to be part of such an amazing work environment, and leaving me some time to complete this PhD thesis properly.

Finally, my last words will be for my mum, who supported me indefectibly throughout these seven years, in the good as well as in the hard times. I dedicate this thesis to her, as well as to my father, who undeniably played an important role in my choices and accomplishments, even though he was not there to see them.

Contents

General Introduction **6**
0.1 Motivation . . . 6
0.2 Decision Making under Uncertainty . . . 7
0.2.1 Bounded, Ecological and Adaptive Rationality . . . 7
0.2.2 Heuristics and Cognitive Biases . . . 8
0.2.3 Social Influence and Decision Making . . . 10
0.2.4 The Rise of Misinformation . . . 12
0.3 The Wisdom of Crowds . . . 14
0.3.1 Foundations . . . 14
0.3.2 Minimal Conditions . . . 15
0.3.3 Impact of Social Influence . . . 16
0.4 Swarm Intelligence . . . 17
0.4.1 Fundamental Principles . . . 18
0.4.2 Pedestrian Dynamics . . . 19
0.5 Outline . . . 21

1 Impact of Information Quantity and Quality on Estimation Accuracy **23**
1.1 Introduction . . . 23
1.2 Experimental design . . . 24
1.2.1 First experiment in Japan . . . 25
1.2.2 First experiment in France . . . 30
1.2.3 Second experiment in Japan . . . 36
1.2.4 Second experiment in France . . . 36
1.3 Distribution of Estimates and Prior information . . . 38
1.3.1 Normalization and Aggregation of Individual Estimates . . . 38
1.3.2 Generalized Normal Distributions and Information . . . 41
1.3.3 Laplace Distributions of Individual log-Estimates . . . 43
1.4 Individual Sensitivities to Social Influence . . . 45
1.4.1 Distributions . . . 45
1.4.2 Consistent Differences in the Use of Social Information . . . 47
1.4.3 Distance between Personal Estimate and Social Information . . . 48
1.5 Model . . . 49
1.5.1 Preliminary model . . . 49
1.5.2 Complete Model . . . 51
1.6 Impact of Information Quantity on Group Performance . . . 52
1.6.1 Impact of Social Information on the Distributions' Center and Width . . . 52
1.6.2 Impact of Social Information Use on Collective Accuracy . . . 55
1.7 Impact of Information Quality on Group Performance . . . 56
1.7.1 Incorrect Information can Help a Group Perform Better . . . 57
1.7.2 Incorrect Information and Sensitivity to Social Influence . . . 60
1.8 Conclusion . . . 61

2 Impact of Filtered Information on Human Phase Separation Processes 64
2.1 Introduction . . . 64
2.2 Experimental design . . . 66
2.2.1 Experimental setup . . . 66
2.2.2 Experimental protocol . . . 67
2.3 Data collection and preprocessing . . . 70
2.3.1 Structure of data . . . 71
2.3.2 Data preprocessing: synchronization and time rectification . . . 71
2.4 Dynamics of Pedestrian Motion . . . 74
2.4.1 Equations of Motion . . . 74
2.4.2 Interaction Forces . . . 74
2.4.3 Correlation Functions . . . 78
2.4.4 Distributions . . . 80
2.5 Phase Separation under Controlled Filtered Information . . . 82
2.5.1 Model Extension . . . 82
2.5.2 Time Evolution of the Fraction of Beeps and Number of Clusters . . . 82
2.5.3 Impact of Information Quantity on Phase Separation Time . . . 84
2.5.4 Impact of Information Quantity on Phase Separation Quality . . . 85
2.5.5 Model Predictions for Different Group Sizes . . . 86
2.5.6 Phase Separation in more Complex Environments . . . 88
2.6 Conclusion . . . 89

General Discussion 91
3.1 Summary of Main Results . . . 91
3.1.1 Estimation Accuracy in Human Groups . . . 91
3.1.2 Human Phase Separation . . . 94
3.2 Future Research Directions . . . 96
3.2.1 Theory of Social Impact . . . 97
3.2.2 Determinants of Social Information Use . . . 97
3.2.3 Assignment of Social Information . . . 98
3.2.4 Estimates, Opinions and Decision-Making . . . 98
3.3 Concluding Words . . . 99

Bibliography 100

Appendices 109

Résumé Détaillé en Français 116

General Introduction

0.1 Motivation

In a nutshell, this PhD – at the crossroads between Physics of Society, Social Psychology, Behavioral Economics and Collective Behavior – is about (individual and collective) human decision-making, and its relation to information.

The reliability of human judgment has been questioned for a long time, and many studies showed that it is irrational and prone to biases of many sorts [1, 2, 3]. We will discuss it in section 0.2.

On top of these intrinsic limitations to human judgment, modern societies have brought a new, extrinsic type of limitation: *information overload* [4, 5]. Indeed, the soar of globalization and information technologies over the last decades has changed our relationship to information and to others, in several critical ways:

1. Information under its many forms (traditional and alternative media, social networks, advertisement...) invades our lives to such an extent that it becomes increasingly challenging to assess its relevance and reliability.

2. Correlatively, the quantity of information available exceeds by far humans' individual capacities to process and integrate it, which raises the question of the development of information-filtering systems.

3. By interconnecting people and institutions, globalization created interdependencies that have considerably increased the complexities of many problems, and in particular those faced by governments and transnational organizations.

4. The scale of communication has also radically changed, be it in the average number of people one interacts with, or their physical distance from one. Gone are the days when social interactions were mostly limited to relatives and village members.

Such considerations raise concerns about the reliability of individual and collective decision-making. In particular, in a context where an increasing number of problems are out of grasp of any single individual, decentralized approaches to decision-making based on *collective intelligence* principles (see reviews in [6, 7, 8]) become more and more attractive, and have drawn a lot of attention in recent years. The benefits of such approaches could be enhanced by the development of information systems aiming to help humans filter and process the colossal amount of information they are continuously subjected to.

The *Wisdom of Crowds* [9] and *Swarm Intelligence* [10] approaches are among the most prominent: the first one assumes that knowledge is widely distributed among individuals in a group and can – if properly aggregated – be leveraged to solve many problems; the second one, inspired by *collective animal behaviour*, postulates that many complex patterns observed at the

level of populations or societies (bird flocks, fish shoals, sheep herds...) emerge from simple local interactions among their members, and thus that the solution to many collective problems may lie in the deciphering of these local interactions.

In sections 0.3 and 0.4, we will outline the main findings in these two broad research areas, and relate them to the two projects presented in this PhD. Both have in common the investigation of the way people perceive, exchange and process information, individually as well as collectively. To achieve this, we developed original methodologies to precisely control the amount of information received and shared by individuals in our experiments.

We will finish this introduction by briefly outlining the questions addressed in this thesis and the methods employed to answer them, in section 0.5.

In this introduction we don't pretend to provide an exhaustive overview of the research domains explored, but rather to offer a synthetic view of the main concepts and results on which our work is grounded. We provide all references that we judged important, so that the reader can go explore further as s/he wishes.

Our hope is that after reading this introduction, the reader feels comfortable with the notions described, and has a global understanding of our main interests in this research.

0.2 Decision Making under Uncertainty

0.2.1 Bounded, Ecological and Adaptive Rationality

Standard economic theories (rational choice theory) consider perfectly *rational agents*, in the sense that in a decision-making process (*i.e.* choosing among alternatives), they have full information about the alternatives, are able to process this information flawlessly (unlimited computational power), to evaluate the potential outcomes (in terms of costs and benefits) of all alternatives, to compare and rank them according to their own preference pattern ("utility" maximization), and finally make a decision that is consistent (no uncertainty, no bias in judgement) with their computed preference ranking[1] [11, 12, 13].

Human rationality so defined may appear unrealistic in regard of everyday life experiences and a bit of introspection. In particular:

- the information available is often limited and potentially unreliable
- the time allotted to make a decision is almost by definition limited
- the brain itself has limited computational power

Such considerations lead Herbert Simon to elaborate the concept of *bounded rationality*, which takes into account extrinsic as well as intrinsic limitations of the human mind [1]. Because of these limitations, rational agents can't make *optimal* decisions, and rather make "*satisficing*" decisions, *i.e.* decisions that are sufficient and satisfying enough given time, cognitive and environmental constraints [14] (see Figure 1).

Consequently, the rationality of an action cannot be judged according to absolute, objective rules, as in standard rationality theory. Rather, it depends on the subjective goals of an agent

[1] More evolved versions of rational choice theory have been developed, such as *expected utility theory* and *rank dependent utility theory* which take uncertainty ("risk") in choices outcomes into account, but that doesn't change any of the conclusions drawn in this section.

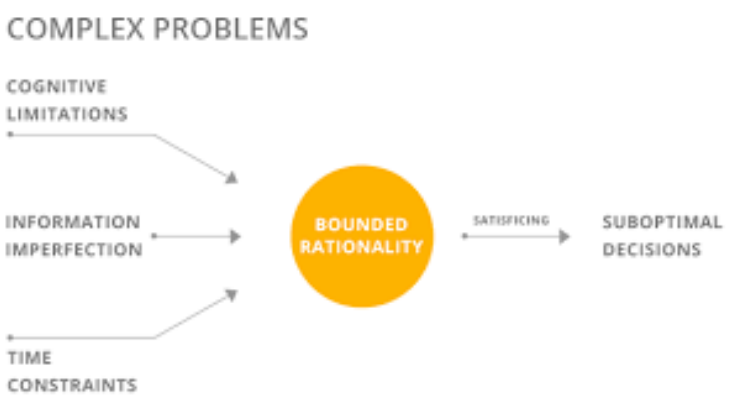

Figure 1: Illustration of the concept of bounded rationality. Due to cognitive limitations, time constraints and limited information, decisions are not made in search for optimality, but for "satisfycing".

in relation to its environment [14]. In other words, rationality is *ecological* [15, 16] and *adaptive* [17, 18].

This new conception of rationality (bounded, ecological and adaptive) is at the core of *behavioral economics*, a branch of economics that tries to incorporate findings from psychology and neuroscience to provide more accurate descriptions of human decision-making. The most prominent advance in this field came from Daniel Kahneman and Amos Tversky's works on decision making under risk and uncertainty, which we will discuss in the next section.

0.2.2 Heuristics and Cognitive Biases

Once acknowledged the above described limitations of the human mind, remains to understand the strategies and mental processes involved in decision making. If in most situations involving a choice, people cannot assess all alternatives and related outcomes by lack of time, information or cognitive abilities, then they need to use rules of thumb (mental "shortcuts") to make efficient decisions under such constraints: the *heuristics*.

In the early 1970's, Tversky and Kahneman found that many observed biases in human judgement – systematic misevaluations of probabilities of events – could be explained by one of three simple heuristics [19]:

- *Representativeness*: one assesses the probability that an event E belongs to a category C by the degree to which E matches one's internal representation (stereotype) of C.
- *Availability*: one compares likelihoods of events according to the ease with which information related to these events comes to mind.
- *Anchoring and adjustment*: one has a tendency to adjust estimates of probabilities toward "anchors", which act as starting points to which one's estimates are relative.

These findings lead Tversky and Kahneman to elaborate their *Prospect Theory* in 1979, a model of decision-making under risk concurrent to the classical expected utility theory [20] (a "prospect" here is a couple {value, probability} for a given alternative). In their model, agents weigh alternatives in terms of their subjective value – evaluated in terms of gains and losses with respect to a "reference point" – and probability of occurrence. The main originality of this model is that it incorporates cognitive biases in its very assumptions:

- *Loss aversion*: people have a tendency to be more affected in a negative way by losses than in a positive way by equivalent gains (see left panel in Figure 2).
- *Marginal effects*: a same change in magnitude of gain (resp. loss) has a lesser impact as the absolute value of the gain (resp. loss) increases (see left panel in Figure 2).
- *Isolation effect*: people tend to set similar probabilities as a *reference point* – and thus disregard them – from which they assess gains and losses for other alternatives.
- *Non-linear probability weighting*: people tend to overestimate low probabilities and to underestimate high probabilities (see right panel in Figure 2).

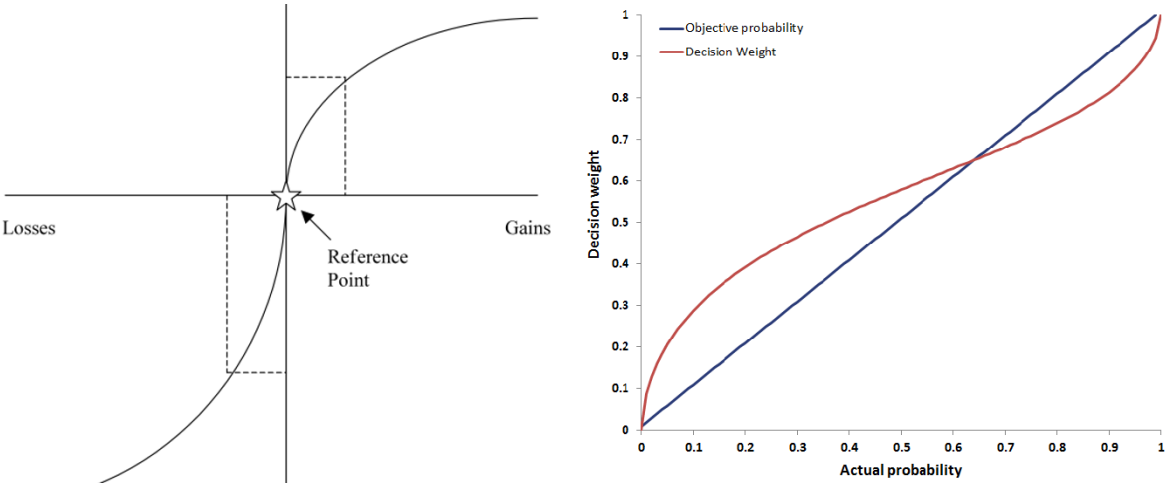

Figure 2: Left: A prospect theory value function illustrating loss aversion marginal effects (concavity for gains, convexity for losses) and loss aversion (steeper curve for loss than for gain); Right: A prospect theory weighting function (in red) illustrating the tendency to overweight low probabilities and underweight high probabilities (inverse-S shape).

Following the vast research program "Heuristics and biases" – launched by Kahneman and Tversky in the early 1970s – many more heuristics have been identified, and with them the origin of many cognitive biases. However, this identification of heuristics with the origin of systematic judgemental errors has lead to interpret heuristics as mere surrogates of optimal and rational statistical procedures [21, 22].

This negative interpretation of heuristics has been strongly criticized, in particular by Gerd Gigerenzer, who argued that heuristics should rather be understood as *evolved* (together with complex psychological mechanisms) strategies (*adaptive toolboxes*), efficient under specific internal (*psychological plausibility*) and environmental (*ecological rationality*) constraints [23, 15]. Thus, different heuristics have evolved to answer different kinds of problems (*domain specificity*), and it can happen that the same heuristics that are efficient in certain situations lead to cognitive biases in other particular situations. Yet, identifying these biases doesn't make heuristics irrational *per se*.

Gigerenzer and Goldstein go even further and argue that simple (*fast and frugal*) heuristics can be not only *psychologically* and *ecologically* rational, but also more efficient than complicated "rational" inference procedures – such as linear multiple regressions or Bayesian models – that take into account all information available and full integration of it [24]. Here are two classical examples of such heuristics, little demanding on knowledge, time and processing power, that have proven very efficient in real-world situations (such as identifying which of a pair of cities is largest):

- *Recognition heuristic*: if of two alternatives to discriminate (according to some criterion), one is recognized and the other is not, choose the recognized one. This heuristic can lead to *less-is-more effects*, when having less prior knowledge about the alternatives entails better inference of the correct one [15].

- *Take-the-best heuristic*: when both objects (or more) are recognized, consider the first available cue (ordered by *cue validity*) able to discriminate between them, and choose the best alternative according to the cue [25].

Many more examples of heuristics and cognitive biases can be found in Ariely's 2008 book "Predictably Irrational: The Hidden Forces That Shape Our Decisions" [3], in Kahneman's 2011 book "Thinking: Fast and Slow" [2], or in the 2016 review by Joyce Ehrlinger [26].

0.2.3 Social Influence and Decision Making

On top of being environmentally and psychologically constrained, our decisions are strongly influenced by others. From the youngest age we experience that others have information that we don't, and that at least partial reliance on social information is necessary to improve our own knowledge. As we grow up and get more confident in our ability to make correct decisions, most of us still intuit that information is widely distributed (*i.e.* that personal information is most of the time incomplete) and that looking at others' decisions or behavior remains a good strategy to gather information and make better decisions. However, the quality of socially transmitted information is often difficult to assess, and reliance on social information can lead to collective behaviors generally considered depreciative, such as *conformity*, *mimetism*, *herd behavior* or *information cascades*. In this section we relate two famous experiments which illustrate these behaviors and more generally highlight the impact of social influence on our decisions.

Asch Conformity Experiments

In the 1950s, Solomon Asch ran experiments – since known as *Asch conformity experiments* – which aimed at studying in a quantitative way the impact of group pressure on decision-making [27]. He developed a methodology that has been used at length since then, which consists in putting a subject in a conflicting situation where his personal opinion (based on strong subjective evidence) is contradicted by other subjects in the same room.

The basic task is simple: a number of subjects (all aside one another in a common room) have to match the length of a line in an image with one of three lines in another image (see Figure 3). All subjects but one (the critical subject) are instructed what to answer beforehand by the experimenter. The task is easy enough (the length of the lines are easily distinguishable) for the critical subject to *know* the correct answer unambiguously. Yet, before giving her opinion (the critical subject is always last to answer), she listens to the answers of other individuals in the room, who all give the wrong answer (as instructed). Individual interviews of critical subjects were carried out posterior to the experiment, with specific questions aiming at understanding the psychological processes at play in subjective choices to yield to the majority or not.

Variants were tried, where the number of instructed subjects varied, the discernibility of the lines decreased, or another naive subject was introduced in order to compare the impact of a high majority with a unanimous majority. The results were probing:

- a large tendency to either adopt or move toward the (wrong) unanimous majority was observed (about one third of estimates), even when the task was easy enough for control groups to be close to 0% wrong answers.

Figure 3: Asch conformity experiment. Left: one subject is naive, while others have been instructed what answers to provide; Right: the task consisted in matching the length of the line in the left image with one of the three lines in the right image.

- the impact of unanimity was huge compared to simple (even high) majorities: a unanimous majority of three individuals had a stronger impact on a subject's decision making than a majority of eight with one dissenter.
- correlatively, the same result highlighted the importance of social support in decision making.
- the impact of a majority decision increased with its size, up to a quite small threshold: majorities of 16 individuals didn't entail a higher rate of yielding than majorities of three.
- the impact of a majority decision increased with diminishing clarity of the stimulus: as distinguishing the length of the lines became harder, the critical subjects became more receptive to social information.
- important individual differences were observed, suggesting a large heterogeneity in humans' sensitivity to social influence.

The author identified three main psychological reasons ("distortions") for yielding to the majority:

- *Distortion of perception*: the social impact was strong enough to make subjects actually *see* the wrong answer proposed by the majority as correct.
- *Distortion of judgement*: subjects still perceived the correct answer, but thought their own perception was incorrect.
- *Distortion of action*: subjects knew they were right, but were afraid of appearing different (in a negative way) to others.

On the other hand, resisting the influence of the majority was mostly due to *confidence* in one's perception or an urge for individuality (resistance to herding).

These findings underline our propensity to give credit to the opinion of others, even when it is obviously absurd.

Drawing Power of Crowds

Stanley Milgram, a famous social psychologist mostly known for his works on *obedience* [28] and the *small world* phenomenon [29], also ran a very insightful experiment on social impact and

herding phenomena [30].

The basic idea of the study was to look at the impact of the size of a stimulus group (crowd) performing a certain act, on the probability that passersby would mimic the group. In the original setup, the stimulus crowd was standing in a busy street of New York City and looked up at a building's window (see Figure 4). It was composed of one to fifteen persons.

Figure 4: Illustration of the "drawing power of crowds": the larger the size of a group of individuals looking up at a building's window (where nothing actually happens), the higher the probability that passersby will stop to look up in turn.

The results indicated unmistakably that larger crowds entailed a greater number of passersby to look up and even *stop* (in order to look up), even though nothing interesting happened at the window. Thus, the probability to herd increases with the number of individuals involved in an activity.

This conclusion sheds light on the concept of *information cascade*: consider a stimulus crowd of five individuals. According to Milgram's results (*i.e.* in this particular setup), about 15 % of passersby would join. But then, the "new" stimulus crowd is soon composed of seven, eight or ten individuals, which will entice even more passersby to join in, which in turn will entail higher stimulus groups sizes, and so on.

It is important to stress that such *herding behavior* has rational foundations. Indeed, in the absence (or lack) of personal information, it is rational to observe the behavior of others, and also rational to assume that statistically more accurate information can be obtained from larger groups of individuals (if something really important or interesting was happening – like a spaceship appearing in the sky – one would like to be aware of it!). Yet, *mimetism* can lead individually rational behaviors to collective "irrational" (or counter productive) outcomes, as in economic crises [31, 32].

0.2.4 The Rise of Misinformation

We have seen that decisions usually rely on both environmental and social information, which are (and always have been) potentially misleading. However, the recent advent and generalization of Internet, and the multiplication of information sources (social networks, blogs, Youtube, alternative media) have radically changed the (specifically) human *informational environment*, thus affecting our relation to information and in turn the way we make decisions. In particular,

it opened the door for *misinformation* to spread at unprecedented levels.

Most *information behavior models* in humans[2] consider the information to be sought, processed and exchanged as full and reliable [34]. However, recent fake news scandals in social media contradict such an assumption [35]. The multiplication of information sources and customized content favored the development of *echo chambers* and *filter bubbles*, in which people are at the same time more susceptible to endorse false information if it confirms their pre-existing beliefs and reluctant to corrections that would contradict these same beliefs [36].

The spread of misinformation in online social media is so pervasive that it has been listed by the World Economic Forum as one of the main threats to our society [37]. Interest in this issue has triggered research for solutions, such as rumor detection in online microblogs [38], evaluation of tweets' credibility in real-time [39] or tracking political abuse in social media [40]. Even companies such as Google and Facebook have started tackling the problem, the first one adding a trustworthiness score in the ranking algorithm [41], the second the possibility for users to signal inappropriate content. Yet, none of these solutions can claim full protection of users against false information.

Moreover, a potential control over information raises ethical questions such as who should be the judge of the legitimacy of information? It seems quite obvious that big companies should not be in charge of such decisions, but should the State be? A panel of experts? Voters? To avoid a top-down control of information by an authority, it seems appropriate to develop decentralized approaches to determine the trustworthiness of information. To help built the infrastructure that would underlie such approaches, a deeper understanding of the impact of false information on decision making is necessary.

Ground breaking experiments have shown that memory can be strongly affected by false information (*misinformation effect*) [42], and that truth is less important in people's adherence to information than their own belief system (*confirmation bias*) or the social norms and pressure they are subjected to [37, 43]. Recent findings have also demonstrated that fake news spread faster and deeper (affect more people) than true information on Twitter, especially when they carry political content [44]. Such conclusions underline the potential dangers of misinformation.

However, we challenge the generally assumed discontinuity between beneficial accurate information and harmful inaccurate information. Indeed, we will argue in Chapter 1 that the harmfulness of some incorrect information is measured by *the degree to which* it is misleading, on a continuous scale. As Fox says, misinformation is but a "species" of information, and as such can still be informative albeit false [45]. We will even show that misinformation can, under specific conditions discussed in Chapter 1, be used as a *nudge* (see definition and examples of nudges in [46]) to boost collective decision making.

In this section, we have learned that rationality is bounded by at least three main factors: limited amount of information available and limited time and cognitive abilities to process it. To overcome these limitations, living beings and humans in particular use evolved strategies called heuristics to make efficient decisions in specifics contexts (psychological capacities and environmental constraints). Heuristics can also be ill-adapted in certain situations and lead to systematic biases in judgement. We have also seen that social information is a crucial component of decision making: mimetic strategies often prove useful and sometimes even necessary,

[2] In Wilson's sense: "By information behavior is meant those activities a person may engage in when identifying his or her own needs for information, searching for such information in any way, and using or transferring that information." [33]

but can also produce unwanted collective outcomes such as conformity and information cascades. Finally, we highlighted the urge to understand the processes at play in the use of information in the light of the recent soar of misinformation related problems.

In the next section, which lays the foundations for Chapter 1, we will focus on a particular type of decision making: the estimation of quantities. We will see that estimation tasks are a very convenient framework to study human decisions in a quantitative way.

0.3 The Wisdom of Crowds

The Wisdom of Crowds consists in tapping into the distributed knowledge of a collection of individuals in order to estimate a value or the probability of an event (see Figure 5). It is thus a particular form of *collective intelligence*, in which individuals in a group don't really "solve" problems (estimations) by themselves. Indeed, "tapping" into the distributed knowledge amounts to aggregate properly the different opinions, which doesn't come *from within* the group.

In its basic form, individuals of a group don't interact with each other, and there has been a long debate about the beneficial or detrimental effect of social interactions on the Wisdom of Crowds. In this section, we will discuss this successful approach to problem solving, its conditions of applicability and the impact of social interactions on it.

Figure 5: Illustration of the concept of Wisdom of Crowds: knowledge is distributed, such that crowds can be, in specific conditions, "smarter" than individuals.

0.3.1 Foundations

In his famous 1907 experiment (see Figure 6), Francis Galton went to a weight-judging competition in Plymouth, where a price would be discerned to the best estimates of the weight of an ox [47]. He gathered all participants' estimates, and showed that the middlemost estimate (*i.e.* the median) fell within 1% range of the real value, *better than the best estimates in the group* [48]. He called this effect the "vox populi", popularized in 2005 as "Wisdom of Crowds" by James Surowiecki [9].

The Wisdom of Crowds (WOC) has drawn a lot of attention since then, and has proven effective in very different contexts, such as political, sport, economic or weather forecasts, or

even medical diagnosis [49, 50, 51, 52, 53, 54, 55]. The WOC is based on the fair assumption (related to our discussions on limited rationality in section 0.2) that **knowledge is distributed, and thus that collectives hold more information than individuals**.

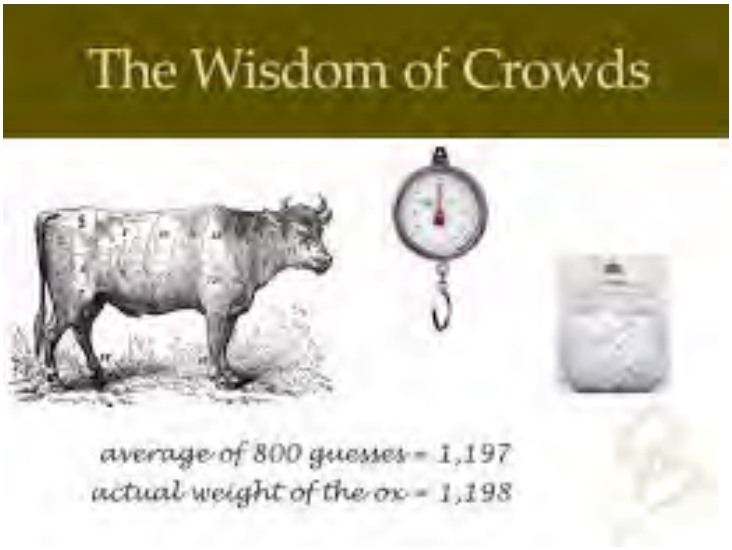

Figure 6: Illustration of Galton's original experiment, where participants in a fair had to estimate the weight of an ox. The median estimate fell impressively close to the exact answer, closer than any participant in the fair.

From this assumption result the following questions:

1. In a given situation where a decision has to be made, how do we find the information holders?
2. How do we best aggregate their individual information into a collective knowledge?
3. How do we infer the truth from this aggregated knowledge?

Answering these questions is the main focus of the WOC research community. The most common way – and the one we will follow – to tackle the problem is to walk on Galton's footsteps and use estimation tasks. Below we will discuss the minimal conditions that have been identified for the WOC to be applicable.

0.3.2 Minimal Conditions

For the WOC to work efficiently, the following conditions are generally assumed necessary:

1. A sufficient amount of knowledge is distributed in the population of interest.
2. There exist a proper *aggregate measure* (mean, median, trimmed mean...) to leverage this distributed knowledge into a good approximation of the truth.
3. Estimates must be *numerous* enough for this approximation to be significant.
4. Estimates must also be *diverse* enough for underestimates and overestimates to cancel out in the aggregation process.
5. Since humans have a tendency to herd (see section 0.2.3), opinions must remain *independent* in order to maintain diversity. The Condorcet's Jury Theorem is also often invoked to highlight the importance of the independence condition (although conclusions on binary decisions don't necessarily apply to continuous estimation tasks) [56, 57].

For a certain quantity to estimate, the proper aggregate measure depends on the properties of the experimental distribution of estimates. To illustrate this, consider that in most recent works, distributions of estimates have been found highly right-skewed [58, 59, 60, 61], which obviously makes the mean an improper measure of the collective wisdom. The median could therefore be a more stable alternative, but a more thorough investigation showed that these right-skewed distributions are actually log-normal (the distribution of the logarithm of estimates is normal), and often centered pretty close to the truth [58, 59]. For such distributions, a proper aggregation method (better than the median) is to take the mean logarithm of estimates μ, which is a good estimator of the center of the distribution, and then approximate the truth by 10^{μ}. Normalization and aggregation of estimates will be discussed at length in section 1.3.1 of Chapter 1.

The *diversity* and *independence* conditions have been challenged in the light of the ubiquity of social interactions. Indeed, in real life situations, people influence each other, which makes opinions rarely – if ever – independent. Even in simplified versions of reality (as in experimental setups for estimation tasks) where people are not allowed to interact directly with each other, their pseudo-independent judgements are part of a shared culture, and thus results of past, often mediated, interactions. These considerations have entailed a long (and still ongoing) debate to know whether, and in what conditions, social interactions impair the WOC or not. In the next section we review and discuss some of the main contributions to this debate.

0.3.3 Impact of Social Influence

In the framework of estimation tasks, *social interactions* are most of the time (computer) mediated: after providing their personal estimate, subjects are given (social) information about the estimates of other individuals, of the same group or not. Social information can be presented in different ways: one or several estimates, the average or median of them, as a distribution... The pieces of information themselves can be chosen either randomly or in more specific ways. After they have received social information, subjects are free to update their initial estimate, which allows to measure the degree to which they have been *influenced* (*i.e.* the degree to which they have changed their opinion).

In their influencial work, Lorenz et al. argued that *social influence* is detrimental to the WOC, because it diminishes diversity – defined as the variance of log-estimates – without reducing the "collective error", defined as the squared deviation of the aggregate measure (in their paper the mean log-estimate) from the truth [58].

In terms of distributions, this result means that the distribution of estimates has narrowed around its center after social influence, but that the center has not moved. Using this fact, the authors argue that this narrowing would increase the confidence of an external assessor on an erroneous value (the center of the distribution, which is only an approximation of the truth), and therefore worsens the collective judgement. However, we challenge this view:

- the argument of the external assessors only holds if their initial opinion is *better* than the aggregate estimate of the crowd. If it is worse (which is likely, as searching for social information arguably means that personal information is scarce), then trusting the aggregate estimate more will get them closer to the truth, even if they will still be wrong.

- this view also amounts to believe that five values close to the truth and ninety five very far are better from a decision making perspective than one hundred values close to the truth. We don't endorse such a conclusion.

- the *diversity prediction theorem* – which the authors are aware of – proves that a reduction in diversity (defined as the variance of estimates) entails an improvement in the average individual error (defined as the average squared deviation from the truth). In other words, estimates get on average closer to the true value after social influence, which we believe to be more an improvement than a deterioration of the collective judgement.

In most papers, the WOC is either defined as the distance from an aggregate measure to the truth (*e.g.* the "collective error" above cited), or as an aggregate measure of individual distances to the truth (*e.g.* the average squared deviation from the truth). For instance, Yaniv showed that exposure to advice improves "judgment accuracy", defined as the mean absolute deviation from the truth [62].

Specific conditions have even been identified under which social influence *improves* the performance of a group:

- *Selective social information*: King et al. found that providing subjects with the best previous estimate in sequential estimation tasks improves their accuracy, in the sense that the median estimate gets closer to the truth after social influence [63];

- *Initial configuration of the population*: Mavrodiev et al. argue that the negative or positive impact of social influence on the WOC is determined by the initial accuracy (derived from the "collective error" as defined in Lorenz's paper) and diversity of estimates (variance of log-estimates) [61].

- *Group discussion*: in another context, Klein et al. showed that group discussion improves collective lie detection ability for untrained groups, compared to simple aggregation of opinions (WOC) [64].

- *Selective tool*: Madirolas et al. remarked that individuals who resist social influence (coined "confident") are on average (geometric mean) more accurate than others, arguably because they are better *informed*. Social influence can thus be used as a selective tool to identify informed individuals and boost the WOC [60].

Although we have seen that social influence can in certain conditions lead groups to herd onto suboptimal behavior, the findings discussed in this section suggest that it can also improve individual and collective decisions, if properly used. In Chapter 1 we will investigate how the *quantity* and *quality* of social information impacts collective accuracy in estimation tasks.

0.4 Swarm Intelligence

Swarm Intelligence is a concept used to describe collective behaviors observed in animal – in particular insect – groups or societies, which complexity seems way beyond the (assumed) simple cognitive abilities of their individual members [10]. Contrary to the Wisdom of Crowds, social interactions are at the heart of Swarm Intelligence, and solutions to problems emerge *from within* groups.

In this section we will describe the fundamental principles underlying Swarm Intelligence and provide typical examples observed in Nature. We will see that these principles also apply to human societies, and in particular to a specific class of human behavior: *pedestrian motion* [65]. The relative simplicity of this behavior offers an ideal framework to study non-verbal social interactions and the collective processing of information [66].

0.4.1 Fundamental Principles

From traffic organization [67] and army foraging [68] in ants to nest buiding in termites [69], very complex collective behaviors have been observed in many animal societies that don't mirror equally complex individual behaviors (see reviews in [70, 71]) – such societies can be seen as *complex adaptive systems* [72]. Key concepts borrowed from *systems theory* help explain the emergence of such complex patterns from simple individual behaviors:

- *Self-organization*: a collection of individuals self-organizes into a particular structure or pattern if they do so in a *decentralized* manner, without external control or the capacity for any individual to "grasp" the global structure [73].

- *Local interactions*: the global structure *emerges* from local interactions among individuals, without being explicitely "coded" in their abilities [73, 74]. For instance, ant colonies are able to discriminate very efficiently among food sources, by a simple system of pheromone laying: foragers go visit sources at random, and on their way back to the nest, the quantity of pheromones they lay on the ground is an increasing function of the nest quality. Then, by simple attraction to pheromones, other foragers are more likely to follow paths where more pheromones were laid. Thus, by reinforcement, all foragers soon go exploit the best food source [75, 76]. The same pheromone-based system allows ants to identify the shortest way to a food source [77].

  Notice that in this example, local interactions are *stigmergic* [78, 79], *i.e.* they are mediated by the environment (pheromones laid on the ground). Social interactions can also be direct, as in the waggle dance of bees: instead of laying pheromones, bees dance in front of their peers in order to indicate the direction, distance and quality of food sources [80, 81].

- *Non-linearity*: these interactions are generally non-linear, *i.e.* a change in stimulus intensity (*e.g.* the quantity of pheromones) doesn't necessarily provoke a proportional change in behavioral response (*e.g.* following the pheromones) [10, 82, 70].

- *Positive feedback*: the examples above mentionned highlight the importance of feedback processes in the emergence of collective patterns: by amplification of small *initial fluctuations*, positive feedbacks can lead the whole system in a certain direction in a short period of time. Milgram's 1969 experiment on the "drawing power of crowds" (see section 0.2.3) is another example of positive feedback in human groups [30].

- *Negative feedback*: a contrario, negative feedbacks can counterbalance excesses or undesired outcomes of positive feedbacks. Evaporation of pheromones is a typical example of negative feedback: when an initially rich food source becomes poor (after substantial exploitation by a colony), foragers start laying less pheromones on their way back from it than on other food sources in the vicinity. Yet, if pheromones didn't evaporate, the path to this newly poor source would still be the one with most pheromone on it, and hence the most attractive. The colony would thus be trapped in a sub-optimal solution to the foraging problem and eventually die. It is only because pheromones evaporate that ants are able to leave a poor food source for a better one [83].

  Another example is the predator-prey co-evolution dynamics (see Lokta-Volterra equations [84, 85, 86]): when preys are a plentiful resource, predators proliferate until preys become rare (negative feedback), which entails a decrease in predator population and thus the re-proliferation of prey. This cycle repeats itself again and again.

We have hereby taken examples from insect societies to illustrate the principles underlying Swarm Intelligence, but many other animal groups exhibit stunning complex behaviors, such as fish schools [87, 88], bird flocks [89, 90] and sheep herds [91, 92] (see Figure 7).

Figure 7: Examples of collective behavior (top panels) and achievements thereof (bottom panels) in animal societies. Topleft panel: bird flock; topright panel: fish school; bottomleft panel: ants exploiting the shortest path to a food source; bottomright panel: termite mounds.

The same principles also apply to human societies, as evidenced by emergent structures of incredible complexity such as languages, cultures, religions, nations or economy, to name just a few. The study of these structures is beyond the scope of this thesis.

However, not all collective behaviors in humans are as complex as those just mentioned. In particular, the patterns observed in groups of pedestrians have raised interest in recent decades, in part because of their relative simplicity, but also because of their proximity to collective dynamics observed in animal societies.

In the next section we will give examples of such collective dynamics in groups of pedestrians, and briefly outline the main models used to describe them (a detailed description will be provided in Chapter 2).

0.4.2 Pedestrian Dynamics

Just as animals, pedestrians are able to self-organize efficiently to respond to specific situations, such as facilitating the flow in busy streets (by forming two lines of opposite walking directions) or at bottlenecks (by alternating the walking circulation), or creating trails in natural environments [66, 93, 94, 95] (see Figure 8).

However, outcomes of collective behaviors in pedestrians are not always positive ("intelligent"), as evidenced by crowd disasters observed sometimes in mass events, when the density of pedestrian becomes very high (sports games, concerts, political meetings, exhibitions...) [95, 96, 97]. Understanding the mechanisms underlying such collective phenomena is thus crucial to assess building or urban layouts and design efficient evacuation systems.

A successful approach in this direction has been to conceptualize pedestrians as *Brownian agents* [98] subjected to *social forces* [99]. Social forces represent the internal "motivation" of a pedestrian to behave in a certain way, and are functions of the information s/he receives from the environment (neighbors and obstacles).

Figure 8: Example of collective patterns in pedestrian crowds. Left: formation of lanes of opposite walking direction; Center: formation of trails in natural environments; right: crowd disasters, as exemplified in this stampede.

In Helbing's first social force model, a pedestrian aims to reach a certain point in a continuous space, and while doing so, s/he continually adjusts to a desired speed, tries to maintain a certain distance to other pedestrians and obstacles (repulsion forces) or wants to join particular people (attraction forces). A stochastic (Brownian) term is introduced into the equations of motion, to represent random fluctuations in behavior due to micro-influences not included in the model [99, 100].

This model was able to fairly reproduce several self-organized processes such as the oscillation of passing direction at bottlenecks, and lead to promising predictions. For example, it has been shown that pedestrian flows can be improved by adding elements in the walking area, such as trees in the middle of large streets or at crossroads, or funnel-shaped constructions at bottlenecks [94].

A generalization of this model, the *Active Walker Model*, has been developed later to include the explicit *action* of pedestrians on their environment (they can change their environment locally). This model explains in particular the stigmergic self-organized phenomenon of trail formation: by laying tracks on deformable grounds, pedestrians increase the probability that other pedestrians will follow the same tracks. Initial fluctuations can thus, by simple reinforcement (see previous section), lead to a massive use of a certain path by further pedestrians. This is very similar to the mechanism of trail formation in ants, also explained by this model [101, 93].

Although these models were originally purely theoretical (trying to reproduce qualitatively field observations), a lot of experimental work has followed to calibrate them and provide more precise forms of the interactions from data [102, 103, 104, 105]. However, some inconsistencies between these models and experimental data, as well as the difficulty to calibrate them, lead researchers to try more heuristic approaches based on the visual information accessible to pedestrians [102]. Such models, although simpler than the classical ones based on Newtonian equations, were also able to reproduce fairly some observed patterns of collective motion.

In Chapter 2, we stick to a Newtonian model, but combine it to a behavioral approach based on the information shared and (mediately) processed by individuals in groups of pedestrians. We also describe a method to extract the functional form of interactions directly from the data.

Neighbors play a central role in our approach, justified by the notions of *locality* (self-organization is based on local interactions) and *bounded rationality*. Indeed, given the size of certain societies (especially insect), and the limited cognitive capacities of their members, it seems likely that individuals rely on local information coming from their *closest neighbors* only (and from the environment of course) to make decisions.

Moreover, the number of neighbors an individual is able to take into account in its decision making should depend on its perceptual and cognitive abilities. The neighborhood is generally defined either as metric or topological:

- *Metric*: individuals are able to consider the information coming from all neighbors within a certain distance (radius) from them. No limit is imposed on the number of neighbors within the circle.
- *Topological*: individuals are able to consider the information coming from a fixed number of neighbors, whatever their distance.

Topological models are generally better at explaining phenomena such as fast density changes in bird flocks or fish schools, because they allow a stronger cohesion in case of perturbations in interindividual distances, such as under a predator's attack [106]. We will also use a topological definition of neighborhood in Chapter 2.

0.5 Outline

This thesis is composed of two main chapters, which common guideline is the cognitive and behavioral processes at play in individual and collective decision making in human groups. We have chosen to tackle this issue through two particular frameworks that are both convenient for quantitative studies and well adapted to our information-based approach (see below): the *Wisdom of Crowds* and the *Swarm Intelligence*.

In this introduction, we have reviewed the main findings and described the most important concepts in both these research areas. In Chapters 1 and 2, we bring our contribution to both domains by taking on the problem from a new perspective, *i.e.* with an emphasis on the *information received and exchanged by members of a group*.

More specifically, we used two very different types of tasks, and introduced new methodologies to rigorously control the amount of information available to individuals:

- In estimation tasks, we controlled the *quantity* and *quality* of information exchanged by individuals through a system of dummy "informers", and looked at the results in terms of individual and collective estimation accuracy;
- In segregation tasks, we designed an artificial sensory interface aiming at providing walking pedestrians with *filtered* information about their environment, and looked at how a controlled amount of information processed by this interface would affect their collective behavior.

We now briefly describe these tasks and methods, and outline the issues discussed in this thesis.

Collective Estimation Tasks

In the Wisdom of Crowds related project, subjects had to estimate quantities of various sorts, thus providing their personal estimates. For each quantity, they received as social information the average estimate of previous participants, and could then revise their initial estimate. We defined their *sensitivity to social influence* as the extent to which they revised their opinion in favor of social information.

We designed an original procedure to control the information received by the subjects, unbeknownst to them: in the computation of the average estimate that served as social information, we introduced dummy estimates, the number (*quantity*) and value (*quality*) of which we controlled. To strengthen the effect of social information, we chose very "hard" (low demonstrability) quantities to estimate (in contrast to most previous studies), in the sense that we expected subjects to have very little prior information about them. This controlled setup allowed us to precisely measure the individual and collective responses to the quantity and quality of the information available and socially exchanged.

We found that distributions of estimates for such "hard" quantities deviated from the usually reported log-normality, which has important consequences on the proper way to aggregate estimates. We provide an interpretation of distributions in terms of the amount of prior information subjects have about a quantity. We also show that individual responses to social information are highly heterogeneous, and depend on the distance between social information and personal opinion. We discuss the efficiency of different strategies (in the way to use social information) in regard to estimation accuracy.

Spatial Segregation Tasks

The second type of task involved pedestrians walking in circular arenas, who had to segregate into groups of the same "color" (they were randomly assigned one of two colors prior to each experiment). They didn't know which group they were part of (*i.e.* their color), and didn't have visual access to the color of other subjects.

To perform the task, they could only rely on acoustic cues ("beeps") coming from electronic devices (tags) attached to their shoulders. These tags also delivered in real time the position (determined by sensors attached on the walls) and color of all subjects in the arena to a central server. The whole system (server + tags + sensors) acted as an artificial *sensory device* (like the retina) able to filter and process complex information in input (positions and colors of all other pedestrians) and return a bit of information (beep or no beep) in output. A pedestrian's tag would beep under specific conditions allowing us to control the amount of information used.

Similarly to the other project, subjects were not aware of this control, and were simply told that they would beep if their "environment" was of the other color. The subjects always received the same kind of information (a beep or no beep), but the conditions for beeping changed. We thus looked at how variations in the amount of information used would impact the segregation time and quality, defined as the number of clusters at final time by analogy with phase separation processes.

This project was also the opportunity to develop a Newtonian model of pedestrian motion, where the functional form of the interaction forces (with the borders of the arena and between pedestrians) were directly "extracted" from the data. To do so, we applied a recently developed methodology that has proven successful in fish [107].

This introduction aimed at providing the reader with a global vision of the issues addressed in this thesis, what motivated them and the methods employed to tackle them. We also wanted to get her/him acquainted with the major concepts and former results on which our research is based, and that we will use all along this thesis.

Chapter 1

Impact of Information Quantity and Quality on Estimation Accuracy

1.1 Introduction

In a globalized, connected, and data-driven world, people rely increasingly on online services to fulfill their needs. AirBnB, Amazon, Ebay, or Trip Advisor, to name just a few, have in common the use of feedback and reputation mechanisms [108] to rate their products, services, sellers, and customers. Ideas and opinions increasingly propagate through social networks such as Facebook or Twitter [109, 110, 111], to the point that they have the power to cause political shifts [112].

Moreover, the digital revolution has led to an exponential increase in the number of media sources and the amount of information they generate and deliver to the population. Yet paradoxically, this information overload has increased the difficulty to verify information, understand an issue or make efficient decisions [4, 5]; up to a certain point, it has also disrupted the relationship between citizens and the truth [36].

Recently the effects of large-scale diffusion of incorrect information on the behavior of crowds have gained increasing interest, because of their major social and political impacts [37]. In particular, there has been recent evidence that fake news propagate faster and deeper (affect more people) than true information on Twitter, especially when they carry political content [44]. In this context, it is crucial to understand how social influence and the diffusion of incorrect information among group members affect individual and collective decision-making.

Two observations can be made about collective decision-making: (a) people usually make decisions not simultaneously but sequentially [113, 114], and (b) decision tasks involve judgmental/subjective aspects. Social psychology research on group decision-making has established that consensual processes vary greatly depending on the demonstrability of answers [115]. When the solution is easy to demonstrate, people often follow the "truth-wins" process, whereas when the demonstrability is low, they are much more susceptible to "majoritarian" social influence [116]. Thus, collective estimation tasks where correct solutions cannot be easily demonstrated are particularly well suited for measuring the impact of social influence on individuals' decisions.

Galton's original work [47] on estimation tasks showed that the median of *independent* estimates of a quantity can be impressively close to its true value. This phenomenon has been popularized as the Wisdom of Crowds (WOC) effect [9] and is generally used to measure a group's performance. Yet, because of the independence condition, it does not consider potential effects of social influence.

In recent years, it has been debated whether social influence is detrimental to the WOC or not: some works argued that it reduces group diversity without improving the collective error [58, 117], while others showed that it is beneficial, if one defines collective performance otherwise [118, 119]. One or two of the following measures are generally used to define performance and diversity: let us note E_i the estimate of individual i, $\langle E_i \rangle$ the average estimate over all individuals, and T the true value of the quantity to estimate. Then, $\mathscr{G}_D = \langle (E_i - \langle E_i \rangle)^2 \rangle$ is a measure of group diversity, and $\mathscr{G} = (\langle E_i \rangle - T)^2$ and $\mathscr{G}' = \langle (E_i - T)^2 \rangle$ are two natural measures of group performance. However, these quantifiers are not independent, since $\mathscr{G}' = \mathscr{G} + \mathscr{G}_D$, which shows that a decrease in diversity $\mathscr{G}_D$ is beneficial to group performance, as measured by $\mathscr{G}'$. Later research showed that social influence helps the group perform better, if one considers only information coming from informed [120], successful [63], or confident [60] individuals (we will see that these traits are actually strongly related). The way social information is defined also matters: providing individuals with the arithmetic or geometric mean of estimates of other individuals has different consequences [60].

Besides these methodological issues, it is difficult to precisely analyze and characterize the impact of social influence on individual estimates without controlling the quality and quantity of information that is exchanged between subjects. Indeed, human groups are often composed of individuals with heterogeneous expertise, so that in a collective estimation task, one cannot rigorously control the quality and quantity of shared social information, and the quantification of individual sensitivity to this information is hence very delicate.

To overcome this problem, we performed experiments in which subjects were asked to estimate quantities about which they had very little prior information (*low demonstrability* of answers), before and after having received social information. The interactions between subjects were *sequential*, while most previous works have used a global kind of interaction, all individuals being provided some information (estimates of other individuals in the group) at the same time [60, 58, 117, 62, 118]. From the individuals' estimates and the social information they received, we were able to deduce their sensitivity to social influence. Moreover, by introducing *virtual "informers"* providing either the true value (hence coined "experts") or some incorrect information in the sequence of estimates – without the subjects being aware of it – we were able to control the *quantity and quality of information* provided to the subjects, and to quantify the resulting impact on group performance.

We discuss the distributions of estimates and their relation to information, as well as normalisation and aggregation methods, and show that subjects' reaction to social influence is heterogeneous and depends on the distance between personal opinion and social information.

We show that social influence helps the group improve its properly defined collective accuracy, and all the more so as more controlled and reliable information is provided (unbeknownst to subjects).

We use the data to build and calibrate a model of collective estimation, to analyze and predict the impact of information quantity and quality received by individuals on the performances at the group level.

Finally, we find that providing a moderate amount of incorrect information to individuals can counterbalance a human cognitive bias to systematically underestimate quantities, and thereby improve collective performance.

1.2 Experimental design

We conducted overall four series of experiments, two at Hokkaido University in Japan, and two at the Toulouse School of Economics in Toulouse, France. The experimental protocol was

similar in the first two experiments, which main focus was the impact of the quantity of information provided to individuals in a group on their individual and collective accuracy. The results obtained in the first one (in Japan) were used to fix the parameters for which we conducted more in-depth investigation in the second one (in France). The third one was a control (and partial) experiment conducted in Japan, in which we asked exactly the same questions as in the first experiment in France. Finally, the fourth experiment (conducted in France), aimed to understand the impact of information quality on individual and collective decision making.

In France, the aims and procedures of the experiments conformed to the Toulouse School of Economics Ethical Rules. All subjects provided written consent for their participation.

In Japan, informed consent was obtained from each participant using a consent form approved by the Institutional Review Board of the Center for Experimental Research in Social Sciences at Hokkaido University.

Below we describe in detail the experimental procedure for each of these experiments.

1.2.1 First experiment in Japan

We recruited 180 participants, all of whom were students from Hokkaido University. They were divided in 5 groups of 36 individuals, and one group per day performed the experiment. Everyday, each group was divided into 6 subgroups of 6 individuals, who participated to the experiment at the same time (what we call a session). Subjects could not participate to more than one session.

After entering the experimental room, subjects were placed in cubicles that prevented interactions between them. Before starting the experiment, they were explained the rules, the payment conditions, the anonymity warranty, and were asked to shut down their mobile phones.

They were then asked to estimate 12 different quantities on a computer (see the list of questions below) and their answering time was limited (25s). If they exceeded the time limit, a warning message would appear in red on the screen. The experimental program was written using the Z-Tree software[1] (see Figure 1.1 for a screenshot of the experiment performed in France in 2016).

Each question (quantity to estimate) involved two steps: first, subjects had to give their personal estimate E_p (note that from now on, the index p will always refer to the "personal" estimates, namely subjects' estimates before they receive social information); then, after receiving the social information I, defined as the arithmetic mean of the τ previous participants' estimates (with $\tau = 1$, 3, or 5), they were asked to give a new estimate E. Subjects were not told the value of τ. They answered the questions sequentially (see Figure 1.2A), the sequence order being that in which they gave their personal estimate.

We devised a mechanism to precisely control the amount of information received and exchanged by individuals: virtual "experts" providing the true value[2] T for each quantity to estimate were inserted at random into the sequence of participants (see Figure 1.2 A). For each sequence involving 36 participants, we controlled the number $n = 0$ or 18 and hence the percentage $\rho = \frac{n}{n+36} = 0\,\%$ or $33\,\%$ of virtual "experts". The social information delivered to human participants, being the average of previous estimates, is hence strongly affected by these virtual "experts". 18 "experts" were inserted in half questions (see list of questions below), at locations chosen randomly prior to the experiment, different for all questions, but the same everyday (for

[1] Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10:171–178.

[2] In the experiment performed in Japan, the value provided by the experts was actually narrowly distributed around the true value (with standard deviation $\sigma_e = 5\%T$).

Figure 1.1: Screeshot of the experience performed in France in 2016 (second round, when the participants receive social information). The first line is the question, the second provides the social information (average estimates of previous participants), the third one allows the participant to make a second estimate, and the last one is where the participants enter their confidence level. On the topright corner is shown the time remaining (in seconds).

all groups). The other half questions had no "experts". The subjects knew nothing about these virtual "experts".

When providing their estimates E_p and E, subjects had to report their confidence level in their answer, on a Likert scale ranging from 1 (very low) to 5 (very high).

At the end of each session, subjects received monetary payments in Japanese Yen (¥) according to their performance, defined as the average closeness to the true value over all questions. Three different intervals for monetary payments were used: ¥2,000 for the best, ¥1,500 for the two next ones, and ¥1,000 for the three last ones.

Remarks:

- An initial condition I_0 was provided to the first subject of each group, as social information. The initial condition was purposely different from the true value, thus creating an artificial bias, to add some difficulty to the collective task. It was chosen from a Gaussian distribution centered around a certain value m_0, with standard deviation σ_0 (note that from now on, the letters m and σ will always refer respectively to the center and width of distributions of estimates). We defined m_0 and σ_0 such that the information thus presented would seem reasonable to the participants (see the list of questions below for the actual values used in the experiments). The second subject in the sequence was provided the average of the initial condition and the first estimate as social information. The third one received the average of the initial condition and the two first answers, and so on until the τ^{th} participant, who was given the average of the τ previous estimates, including the initial condition. After her, all subjects were given the average of the τ previous estimates.

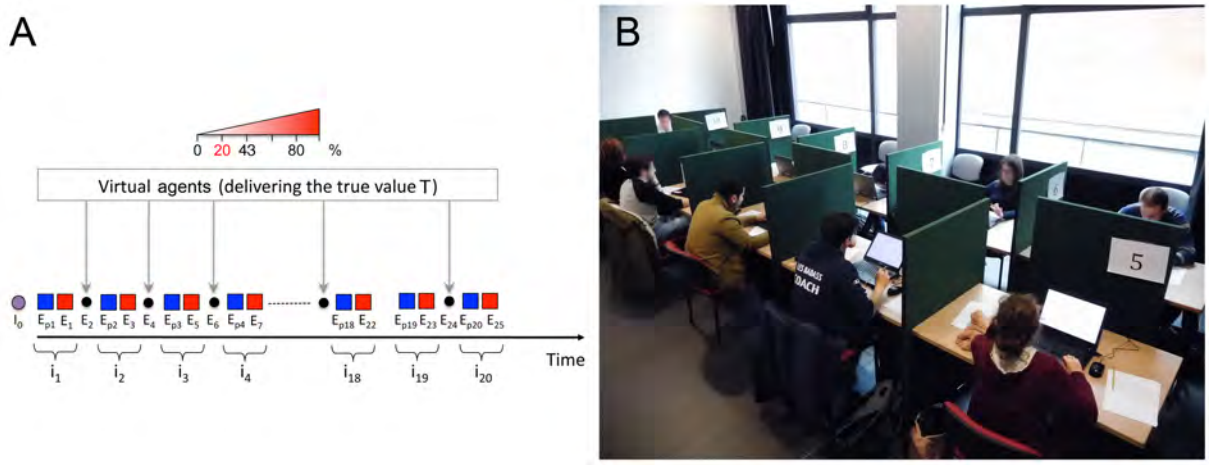

Figure 1.2: The experimental protocol is summarized on the left panel A. For a given question, each individual i_k ($k = 1, ..., 20$) gives first her personal estimate E_p (in blue). Then, after receiving social information (the arithmetic mean of the τ previous estimates in Japan, the geometric mean in France), they give their new estimate (in red). Virtual experts (black dots) are added randomly into the sequence, with percentage ρ ($\rho = 20\%$ in this example, as highlighted in red), without subjects being aware of it. They impact the social information given to the subjects, and can thus be seen as an information input. On the right panel B is a picture of the experimental room (in the first experiment in France).

- As a safeguard against subjects not showing up at a given session, four "extra" subjects were recruited in every session. When too many subjects were there, they participated in another experiment run in parallel of the present one, and unrelated to it.

Questions (asked in Japanese)

1. How old was Gandhi when he died ? ($\tau = 3$ and $\rho = 0\%$)
   - $T = 78$
   - $m_0 = 1.2T$
   - $\sigma_0 = 10\%$
2. How old was Yasujiro Ozu when he died ? ($\tau = 3$ and $\rho = 1/3$)
   - $T = 60$
   - $m_0 = 1.2T$
   - $\sigma_0 = 15\%$
3. Jar 1: How many balls do you think are in this jar ? ($\tau = 1$ and $\rho = 1/3$)

- $T = 450$
- $m_0 = 1.5T$
- $\sigma_0 = 30\%$

4. How many people live in Tokyo (in terms of 10,000) ? ($\tau = 5$ and $\rho = 1/3$)
   - $T = 1,346$
   - $m_0 = 1.25T$
   - $\sigma_0 = 30\%$
5. How many cell phones were sold in Japan in 2014 (in terms of 10,000) ? ($\tau = 1$ and $\rho = 0\%$)
   - $T = 4,000$
   - $m_0 = 0.75T$
   - $\sigma_0 = 30\%$
6. What is the median income per month in Japan, in yen (in terms of 10,000 yens) ? ($\tau = 5$ and $\rho = 0\%$)
   - $T = 36$
   - $m_0 = 0.8T$
   - $\sigma_0 = 20\%$
7. Jar 2: How many balls do you think are in this jar ? ($\tau = 1$ and $\rho = 1/3$)

- $T = 100$
- $m_0 = 1.5T$
- $\sigma_0 = 30\%$

8. How many matches do you think are in this picture ? ($\tau = 1$ and $\rho = 0\%$)

- $T = 400$
- $m_0 = 0.75T$
- $\sigma_0 = 30\%$

9. What is the average distance between Earth and the Moon (in terms of 1,000 km)? ($\tau = 5$ and $\rho = 1/3$)

- $T = 384$
- $m_0 = 1.25T$
- $\sigma_0 = 30\%$

10. What is the average life expectancy for men in Ethiopia ? ($\tau = 3$ and $\rho = 0\%$)

- $T = 63$
- $m_0 = 1.25T$
- $\sigma_0 = 10\%$

11. How many books does the American Congress Library hold (in terms of 10,000) ? ($\tau = 5$ and $\rho = 0\%$)

- $T = 2,389$
- $m_0 = 0.8T$
- $\sigma_0 = 20\%$

12. What is the distance between Tokyo and Pyongyang ? ($\tau = 3$ and $\rho = 1/3$)

- $T = 1,287$
- $m_0 = 1.25T$
- $\sigma_0 = 10\%$

1.2.2 First experiment in France

186 subjects were recruited, the large majority of whom were students from the University of Toulouse. The experimental procedure was very similar to the one described above, but organized in a slightly different way. 20 sessions were organized over 5 days, in which 8 subjects had to answer 29 questions (4 "families" of 6 questions: population of cities, general knowledge, number of objects in images, astronomical features, plus 5 questions given a special treatment; see below). The answering time was limited to 40s per estimate, after which a blinking text urging the subject to speed up would appear on the screen. In each session, the 8 subjects were part of 8 different sequences (defined as in the experiment performed in Japan) associated with the 8 different couples (ρ, τ).

As in the experiment performed in Japan, each question involved two steps: subjects had to first give their personal estimate, and could then revise it after receiving social information. However, a preliminary analysis of the data from the experiment in Japan taught us that:

- the impact of τ could hardly be observed and that we should therefore put more emphasis on ρ (four values instead of two: $\rho = 0\,\%$; $20\,\%$, $43\,\%$ and $80\,\%$, corresponding respectively to $n = 0$, 5, 15 and 80 "experts") than on τ (two values instead of three: $\tau = 1$ and 3)
- the social information should be defined as the *geometric mean* of the τ previous participants' estimates, rather than the arithmetic mean – since humans think in terms of orders of magnitude (see the discussion on distributions of estimates in section 1.3) [121].
- we could increase the number of questions, and we chose to add "harder questions" (lower demonstrability), as they are more relevant to the purpose of the present study, as explained in the introduction.

.

Each question was sequentially answered by 20 human subjects. As in the experiment performed in Japan, virtual "experts" (providing the true value) were inserted into the sequence at random positions (Figure 1.2A). The 8 conditions (ρ, τ) were randomly associated to the 8 participants for each question. It is important to stress that in this experiment, and contrary to the one performed in Japan, *all questions were asked in all conditions.*

The order of questions 1 to 24 was randomized in all sessions, whereas questions 25 to 29 were always in the same order. These five last questions had a special treatment: individuals in a session were divided into 3 subgroups according to their use of social information across the first 24 questions: "followers", "average" and "confident" subjects (we precisely define these categories in section 1.4.2). Then we looked for consistencies in subjects' use of social information across questions: are there tendencies for some individuals to consider more or less the opinion of others? If yes, how differentiable would these tendencies be?

As in Japan, subjects had to report their confidence in their estimates, both before and after social influence, on the same Likert scale as previously described.

At the end of every session, subjects were paid accordingly to their performance (defined similarly to Japan): the first one was paid 20€, the two next ones 15€ and the four last ones 10€.

Remarks:

- The initial condition I_0 was this time chosen proportional to the true value, with a factor $\kappa = 2$, 5 or 50, depending on the questions (see the list of questions below). We alternated the factor between κ and $\frac{1}{\kappa}$, to control for a possible asymmetry between larger and smaller values.

- We added restrictions to the answers subjects could possibly give. These restrictions were defined up to a factor of $\lambda = 2$, 3, 4 or 7 orders of magnitude smaller or larger than the true value (see list of questions). These values of λ were chosen such that answers beyond these bounds could be safely considered as totally absurd (or a typing error).
- For the same reason as explained above, we recruited 10 subjects per session, instead of 8. Overall, we needed 160 subjects, such that 26 were "extra" subjects (we had 186 in total). They were not part of any sequence, nor associated to any condition (ρ, τ), but were not aware of such difference in treatment. They were provided the value κT or $\frac{T}{\kappa}$ (with probability 50 %; T is the true answer) as social information. We recorded all their answers for the statistics. They were paid 10€ whatever their performance.

Questions (asked in French)

1. What is the population of Tokyo and its agglomeration?
   - $T = 38,000,000$
   - $I_0 = \frac{T}{5}$
   - $\lambda = 10^3$
2. What is the population of Shanghai and its agglomeration?
   - $T = 25,000,000$
   - $I_0 = 5T$
   - $\lambda = 10^3$
3. What is the population of Seoul and its agglomeration?
   - $T = 26,000,000$
   - $I_0 = \frac{T}{5}$
   - $\lambda = 10^3$
4. What is the population of New-York City and its agglomeration?
   - $T = 21,000,000$
   - $I_0 = 5T$
   - $\lambda = 10^3$
5. What is the population of Madrid and its agglomeration?
   - $T = 6,500,000$
   - $I_0 = \frac{T}{5}$
   - $\lambda = 10^3$
6. What is the population of Melbourne and its agglomeration?
   - $T = 4,500,000$
   - $I_0 = 5T$
   - $\lambda = 10^3$
7. How many ebooks were sold in France in 2014?
   - $T = 5,000,000$

- $I_0 = \frac{T}{5}$
- $\lambda = 10^4$

8. How many books does the American Congress library hold?
   - $T = 23,000,000$
   - $I_0 = 5T$
   - $\lambda = 10^4$
9. How many people die from cancer in the world every year?
   - $T = 15,000,000$
   - $I_0 = \frac{T}{5}$
   - $\lambda = 10^4$
10. How many cell phones are sold in France every year?
    - $T = 22,000,000$
    - $I_0 = 5T$
    - $\lambda = 10^4$
11. How many kilometers does a professional cyclist bike a year?
    - $T = 40,000$
    - $I_0 = \frac{T}{5}$
    - $\lambda = 10^2$
12. How many cars are stolen in France every year?
    - $T = 110,000$
    - $I_0 = 5T$
    - $\lambda = 10^3$
13. Marbles 1: How many marbles can you see in this jar?

    - $T = 100$
    - $I_0 = \frac{T}{2}$
    - $\lambda = 10^2$

14. Marbles 2: How many marbles can you see in this jar?

- $T = 450$
- $I_0 = 2T$
- $\lambda = 10^2$

15. Matches 1: How many matches can you see?

- $T = 240$
- $I_0 = \frac{T}{2}$
- $\lambda = 10^2$

16. Matches 2: How many matches can you see?

- $T = 480$
- $I_0 = 2T$

- $\lambda = 10^2$

17. Rope 1: In your opinion, how long is this rope (in cm)?

- $T = 200$
- $I_0 = \frac{T}{2}$
- $\lambda = 10^2$

18. Rope 2: In your opinion, how long is this rope (in cm)?

- $T = 700$
- $I_0 = 2T$
- $\lambda = 10^2$

19. What is the radius of the Sun (in km)?

- $T = 700,000$
- $I_0 = \frac{T}{50}$
- $\lambda = 10^7$

20. What is the distance between Earth and the Moon (in km)?

- $T = 385,000$
- $I_0 = 50T$
- $\lambda = 10^7$

21. How many stars does the Milky way hold (in million stars)?

- $T = 235,000$
- $I_0 = \frac{T}{50}$

- $\lambda = 10^7$

22. How many billions kilometers is worth a light-year?

    - $T = 9,000$
    - $I_0 = 50T$
    - $\lambda = 10^7$

23. How many galaxies does the visible universe hold (in million galaxies)?

    - $T = 100,000$
    - $I_0 = \frac{T}{50}$
    - $\lambda = 10^7$

24. How many cells are there in the human body (in billion cells)?

    - $T = 100,000$
    - $I_0 = 50T$
    - $\lambda = 10^7$

25. What is the population of Amsterdam and its agglomeration?

    - $T = 1,600,000$
    - $I_0 = 5T$
    - $\lambda = 10^3$

26. What is the annual gross salary of a professional league 1 soccer player in France (in euros)?

    - $T = 540,000$
    - $I_0 = \frac{T}{5}$
    - $\lambda = 10^3$

27. Matches 3: How many matches can you see?

    - $T = 720$
    - $I_0 = 2T$
    - $\lambda = 10^2$

28. What is the total mass of oceans on Earth (in billion tons)?

    - $T = 1,400,000,000$

- $I_0 = \frac{T}{50}$
- $\lambda = 10^7$

29. What is the distance from planet Aramis to its Sun (in km)?

- $T = 1,000,000$
- $I_0 = 5T$
- $\lambda = 10^7$

1.2.3 Second experiment in Japan

An additional experiment was performed in Japan with the same questions as used in the experiment performed in France (asked in Japanese), but subjects only had to provide their personal estimate E_p. The main motivation for this additional experiment was the difference observed in the distributions of individual estimates in the two first experiments (see Figures 1.6A and 1.6B). We hypothesized that such a difference was mainly due to the difference in questions asked. We indeed found that the same questions lead to a very similar distribution of personal estimates (see Figure 1.7).

1.2.4 Second experiment in France

The procedure was very similar to the one explained in 1.2.2, therefore we will mostly focus on the novelties, the main one being that the information provided to the subjects could be different from the truth (in a controlled range). 231 subjects were recruited, mostly from Toulouse University. 20 sessions were organized over 5 days, in which subjects had to answer 32 questions (the first 27 of which were exactly the same as in 1.2.2, the two following were slightly changed, and 3 were added; see the new questions below).

We controlled the information T_I provided by the "informers" through a parameter α (α = -2, -1, -0.5, 0, 0.5, 1, 1.5, 2, 3) defined such that $T_I = T.K^\alpha$, where T is the true value of a quantity under consideration, and thus controlled the quality of the information provided to the subjects (the closer T_I from T, the higher the information quality). K was chosen such that the information provided T_I would seem reasonable to the subjects even in the extreme cases $\alpha = -2$ and $\alpha = 3$, and thus depends on the quantity to estimate (see Table 1.1). The experimental values of α were chosen according to model predictions from the previous experiments. Notice that $\alpha = 0$ when the "informers" provide the true value ($T_I = T$).

Each participant in each session was associated a value of ρ ($\rho = 0\%$ for subject 1, $\rho = 20\,\%$ for subjects 2 to 5 and $\rho = 43\,\%$ for subjects 6 to 9). For subjects 2 to 9 ($\rho \neq 0\%$; no α for $\rho = 0\,\%$), one particular value of α was associated to each question. Table 1.1 shows the values of α associated to each question for each subject. In this organisation, each condition (ρ, α) is repeated 16 times. Also, a value of τ (1 or 3) was associated alternatively to each question, independently of the values of ρ and α. Table 1.1 below shows the conditions for each subject and each question. These conditions were repeated in every session.

Remarks:

- We imposed no restriction on the participants' answers, because previous experiments proved it unnecessary.
- Because the initial condition was of little impact in previous experiments, it was chosen here simply as the value T_I provided by the "informers". In the particular case $\rho = 0\,\%$ (no "informers"), we chose to provide $\alpha = 0$ (the true value). The questions' order was

randomized in each session, and no special treatment was undertaken (no subgroups for questions 25 to 29, which were randomized just as the others).

- At the end of every session, subjects were paid according to their performance: 20€ for the two first ones, 15€ for the four next ones and 10€ for the three last ones.

- As in the previous experiment, we recruited 1 to 3 "extra" subjects per session (51 came in total). They were not part of any sequence, and the social information they were provided was generated by us from a value of α randomly chosen in the interval [-3,3]. They were not aware of any difference in treatment, and were paid 10€ whatever their performance.

			$\rho = 20\,\%$				$\rho = 43\,\%$			
Question	τ	K	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Subj. 9
1	1	1.5	0.5	-0.5	1	-1	0.5	-0.5	1	-1
2	3	1.6	1	-1	0.5	-0.5	1	-1	0.5	-0.5
3	1	1.5	2	-2	3	1.5	2	-2	3	1.5
4	3	1.5	3	1.5	2	-2	3	1.5	2	-2
5	1	1.6	0.5	-0.5	2	-2	0.5	-0.5	2	-2
6	3	1.6	1	-1	3	1.5	1	-1	3	1.5
7	1	2.5	2	-2	1	-1	2	-2	1	-1
8	3	3.3	3	1.5	0.5	-0.5	3	1.5	0.5	-0.5
9	1	2.5	0.5	-0.5	3	1.5	0.5	-0.5	3	1.5
10	3	2.5	1	-1	2	-2	1	-1	2	-2
11	1	2.2	2	-2	0.5	-0.5	2	-2	0.5	-0.5
12	3	2.5	3	1.5	1	-1	3	1.5	1	-1
13	1	1.3	0.5	-0.5	1	-1	0.5	-0.5	1	-1
14	3	1.3	1	-1	0.5	-0.5	1	-1	0.5	-0.5
15	1	1.3	2	-2	3	1.5	2	-2	3	1.5
16	3	1.3	3	1.5	2	-2	3	1.5	2	-2
17	1	1.6	0.5	-0.5	2	-2	0.5	-0.5	2	-2
18	3	1.8	1	-1	3	1.5	1	-1	3	1.5
19	1	2.8	2	-2	1	-1	2	-2	1	-1
20	3	2.9	3	1.5	0.5	-0.5	3	1.5	0.5	-0.5
21	1	10.1	0.5	-0.5	3	1.5	0.5	-0.5	3	1.5
22	3	7.1	1	-1	2	-2	1	-1	2	-2
23	1	21.1	2	-2	0.5	-0.5	2	-2	0.5	-0.5
24	3	18.1	3	1.5	1	-1	3	1.5	1	-1
25	1	1.6	0.5	-0.5	1	-1	0.5	-0.5	1	-1
26	3	1.7	1	-1	0.5	-0.5	1	-1	0.5	-0.5
27	1	1.4	2	-2	3	1.5	2	-2	3	1.5
28	3	21.1	3	1.5	2	-2	3	1.5	2	-2
29	1	8.1	0.5	-0.5	2	-2	0.5	-0.5	2	-2
30	3	8.1	1	-1	3	1.5	1	-1	3	1.5
31	1	5.9	2	-2	1	-1	2	-2	1	-1
32	3	2.2	3	1.5	0.5	-0.5	3	1.5	0.5	-0.5

Table 1.1: Organization of the parameters in each session. Each subject is associated a specific value of ρ for all questions, but the values of τ and α vary across questions. The "Subjects" columns (Subj. i, $i = 2 \ldots 9$) give the values of α for each question.

Questions (asked in French)

28. What is the total mass of oceans on Earth (in thousand billion tons)?
    - $T = 1,400,000$
29. What is the distance from planet Mercury to the Sun (in km)?
    - $T = 57,800,000$
30. What is the total length of the metal threads used in Golden Gate Bridge's braided cables (in km)?
    - $T = 129,000$
31. What is the mass of Kheops pyramid (in tons)?
    - $T = 5,000,000$
32. How much did Burj Khalifa Tower, in Dubaï, cost (in thousand dollars)?
    - $T = 1,500,000$

1.3 Distribution of Estimates and Prior information

Distributions of individual answers in estimation tasks are generally presented as nearly Gaussian [60, 117]. However, we found that for low demonstrability questions at least, Gaussian distributions underestimate by far the probability of answers very far from the truth, and we thus investigated Cauchy and Laplace distributions as possible alternatives. In this section, we compare these distributions with experimental data, find relations between them, and propose a qualitative interpretation of distributions in terms of the degree of prior information a group has about a certain quantity to estimate.

1.3.1 Normalization and Aggregation of Individual Estimates

A major issue in the Wisdom of Crowds research domain is to find a proper way to aggregate the knowledge of various individuals. In the particular case of estimation tasks, aggregating various estimates requires to carefully choose a normalization process, in order to make estimates of different quantities comparable (especially when they differ by orders of magnitude). In this section we quickly review the most common choices of normalization and aggregation, and provide our own insights based on our experimental data.

Usual Normalizations

In particular cases when some quantities to estimate are very close to each other (difference typically smaller than one order of magnitude), estimates E_i (i is the index for individuals) can be combined more or less adequately without being normalized [118, 63, 62]. More generally, when quantities to estimate span one to several orders of magnitude, a normalization is required, and the simplest way to do so is to divide estimates by each respective quantity q's true value T_q [59]:

$$E_{\mathrm{norm}_{i,q}} = \frac{E_{i,q}}{T_q} \tag{1.1}$$

Moreover, previous works have shown that distributions of independent estimates are generally highly right-skewed, while distributions of their common logarithms are much more symmetric (log-normal distributions) [60, 58, 117]. Indeed, people think in terms of orders of magnitude,

especially when large quantities are involved, which makes the logarithmic scale more natural to represent human estimates [121]. A better normalisation is thus [58, 117]:

$$X_{i,q} = \log(\frac{E_{i,q}}{T_q}) \tag{1.2}$$

Moreover, most commonly used distributions (such as Gaussian distributions) are characterized by their center m and width σ, which represent respectively the collective *bias* of a certain group regarding a certain quantity and the *dispersion* of individual estimates around this bias. They should therefore be considered in the normalization process:

$$Z_{i,q} = \frac{X_{i,q} - m_q}{\sigma_q} \tag{1.3}$$

Before discussing the proper way to estimate these parameters, let us remind and clarify here the notations that will be used all along this chapter:

- the letters m and σ respectively refer to the center and width of a distribution of estimates, which also characterize respectively (as explained right above) the collective bias and dispersion of estimates. When we talk about experimental distributions (most of the time), the same letters refer to the *estimators* of the center and width (see next paragraph);
- the subscript p is specifically used to describe the distributions of estimates *before social influence* (*personal* estimates);
- the index q references the questions/quantities.

Aggregation Rules: Estimators of the Center and Width

What is the best way to aggregate estimates in a group? How do we best measure diversity? To answer such questions, the first step is to be able to satisfyingly estimate the center m and width σ of experimental distributions of estimates.

The best estimators of these parameters actually depend on the type of distribution considered. For Gaussian-like distributions, the mean and standard deviation are respectively the maximum likelihood estimators of m and σ, and can be used for the normalization 1.3 [60, 122, 123]. However, for Cauchy-like distributions, the maximum likelihood estimators are respectively the median and half the interquartile range (IQR), while for Laplace-like distributions, they are respectively the median and the average absolute deviation from the median $\langle |X_i - m| \rangle_i$ (which is actually a measure of *diversity*).

Figure 1.3 shows the bias m_p and dispersion σ_p for questions 1 to 27 (identical in the second experiment in Japan and the two experiments in France), computed for Laplace distributions. Both vary significantly with questions, which shows that the normalization 1.3 is relevant. In panels C and D, one clearly identifies four "blocks" of 6 points (questions 1 to 24), corresponding to the the four families of questions. This suggests that σ_p is characteristic of the type of quantity to estimate, and correlates with the demonstrability of a quantity. In particular, the lower the demonstrability of a quantity, the higher σ_p, and vice-versa (compare city populations related quantities – first block – with astronomy related quantities – fourth block).

Notice that the bias is negative for most questions, reminiscent of the human cognitive bias to underestimate quantities, due to their nonlinear internal representation [124]. As we will show later, this phenomenon has strong implications regarding the influence of information provided to the group.

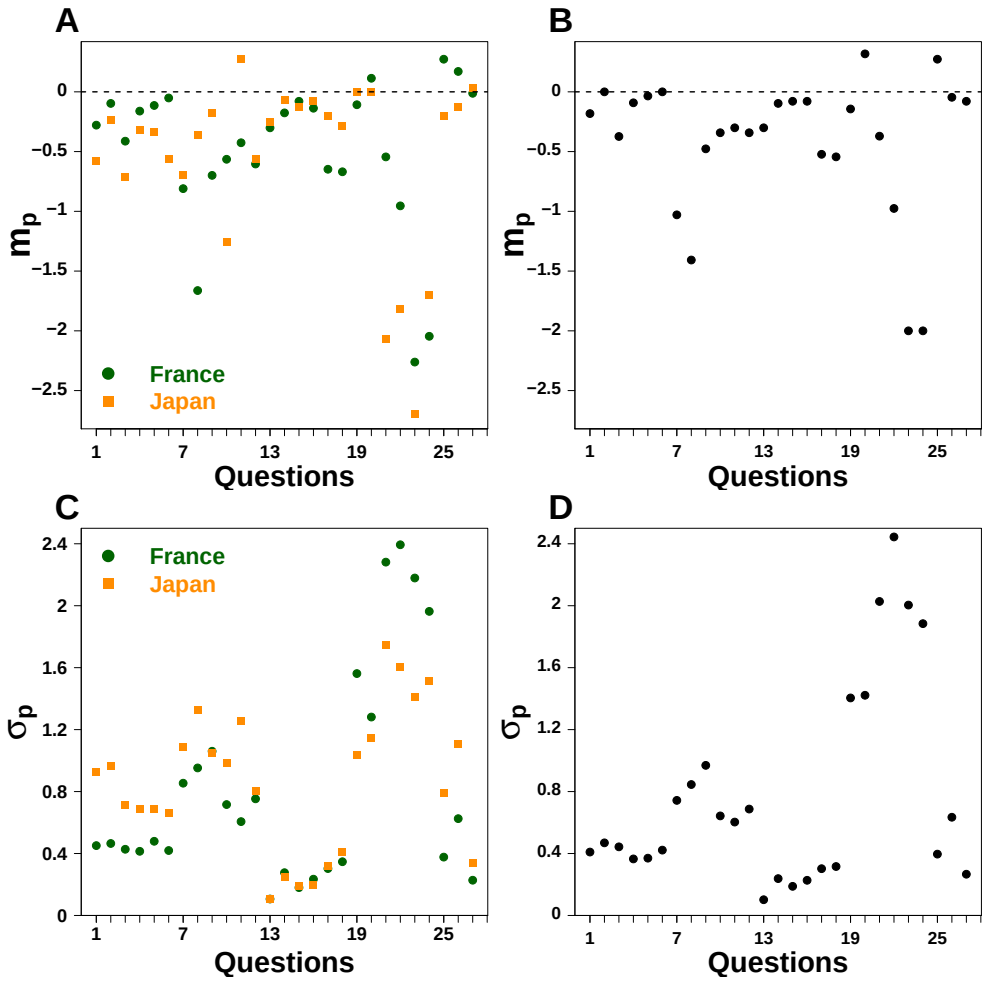

Figure 1.3: Median m_p and average absolute deviation from the median σ_p for the 27 first questions in: A. and C. the first experiment in France (green) and the second experiment in Japan (orange); B. and D. the second experiment in France. Points for questions 28 and 29 are not shown for two reasons: first, they were different in the first experiment in France and second experiment in Japan on one hand, and in the second experiment in France on the other hand; second, because they lead to far outlying values of m_p and σ_p, which would diminish the graph's readability if we included them. Notice that four "blocks" of 6 points (corresponding to the four families of questions from question 1 to 24) can be easily distinguished in panels C and D, suggesting that σ_p is characteristic of the type of quantity to estimate.

In Figure 1.4, we show the distribution of Z for the four experiments we performed, and compare theoretical fits from Cauchy, Laplace and Gaussian distributions. Laplace distributions (exponential decay on both sides) are undoubtedly favored in all cases. The estimators for Laplace distributions (median and average absolute deviation from the median) have been used for the normalization, such that the four distributions collapse onto the standard Laplace distribution (center 0 and width 1). In Figure 1.5, we combine all data from the four experiments.

Notice that since the estimators of the width for Gaussian (standard deviation) and Cauchy (IQR) distributions are different from that of Laplace distributions, we have to compare the standard Laplace not with the standard Gaussian or Cauchy distributions, but with the Gaussian and Cauchy distributions of same width (respectively $\sigma_{\text{Gauss}} \approx 1.1457\,\sigma_{\text{Laplace}}$ and $\sigma_{\text{Cauchy}} = \log_e(2)\,\sigma_{\text{Laplace}}$.

As a consequence, adequate choices of a Wisdom of Crowds indicator (mean or median) and definition of diversity depend on the shape of the experimental distribution found. In the next

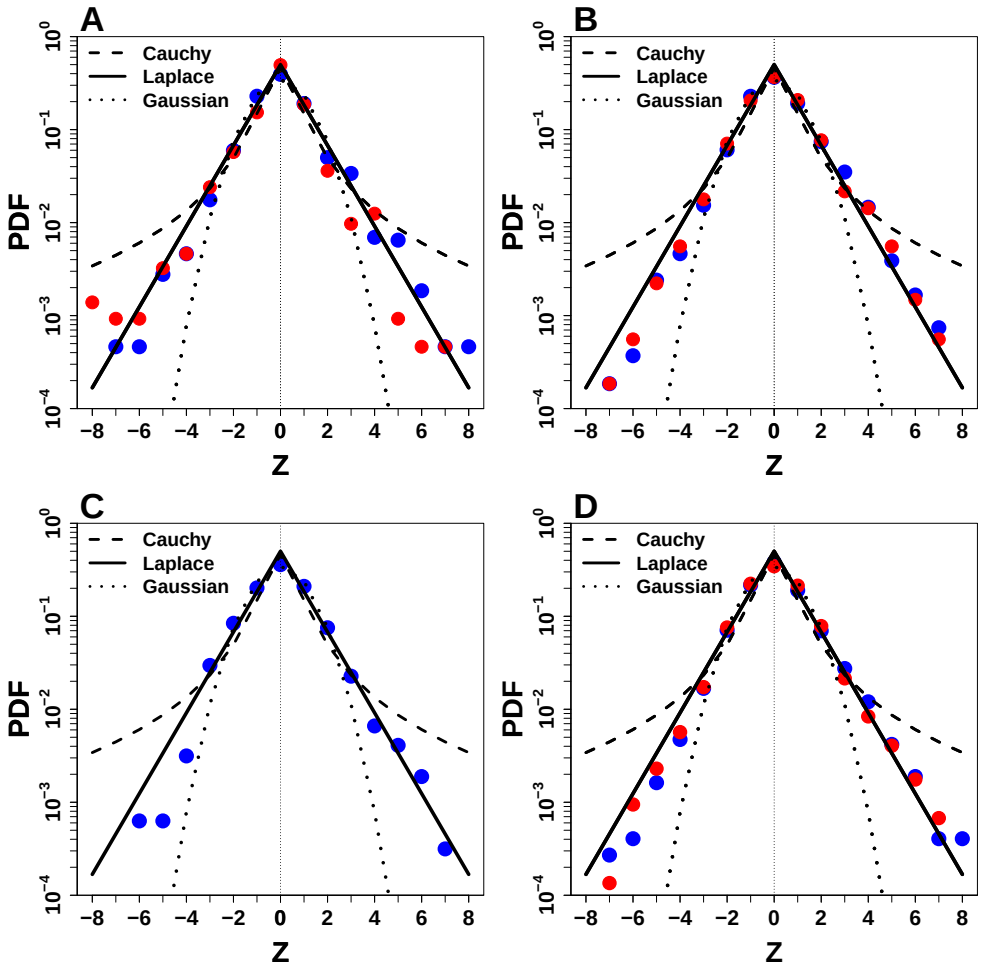

Figure 1.4: Distribution of normalized individual estimates $Z_{i,q} = \frac{X_{i,q} - m_q}{\sigma_q}$ *in A. the first experiment in Japan; B. the first experiment in France; C. the second experiment in Japan; the second experiment in France, before (blue) and after (red) social influence. The black lines are the standard (center 0 and width 1) Laplace distribution (full line), the Cauchy distribution (dashed line) of same width* ($\sigma_{\text{Cauchy}} = \log_e(2)\,\sigma_{\text{Laplace}}$) *and the Gaussian distribution (dotted line) of same width* ($\sigma_{\text{Gauss}} \approx 1.1457\,\sigma_{\text{Laplace}}$). *Laplace distributions fit best the experimental data.*

subsection we discuss why certain studies find experimental distributions of estimates close to Gaussian distributions, while others (like ours) find them closer to Laplace distributions.

1.3.2 Generalized Normal Distributions and Information

In his 1923 fundamental paper, Edwin Bidwell Wilson remarked that in many empirical data sets, the frequency of an error X (deviation from a certain value) follows Laplace's first or second law of error [125], namely can be expressed as an exponential function of respectively the absolute value of the error ($f(X) \propto e^{-\beta|X|}$; Laplace distribution) or of the square of the error ($f(X) \propto e^{-\beta X^2}$; Gaussian distribution) [126]. One year later, Rider proposed a more general law: the frequency of an error can be expressed as an exponential function of the n^{th} power ($n > 0$) of the absolute value of the error ($f(X) \propto e^{-\beta|X|^n}$) [127], encompassing Gaussian and Laplace distributions into a continuous family of exponential distributions, later coined Generalized Normal Distribution (GND) family and studied in detail by Nadarajah in [128]. The PDF of those distributions reads:

$$f(X, m, \sigma, n) = \frac{1}{2\,\sigma\,\Gamma(1 + 1/n)} \exp\left\{ - \left| \frac{X - m}{\sigma} \right|^n \right\}, \tag{1.4}$$

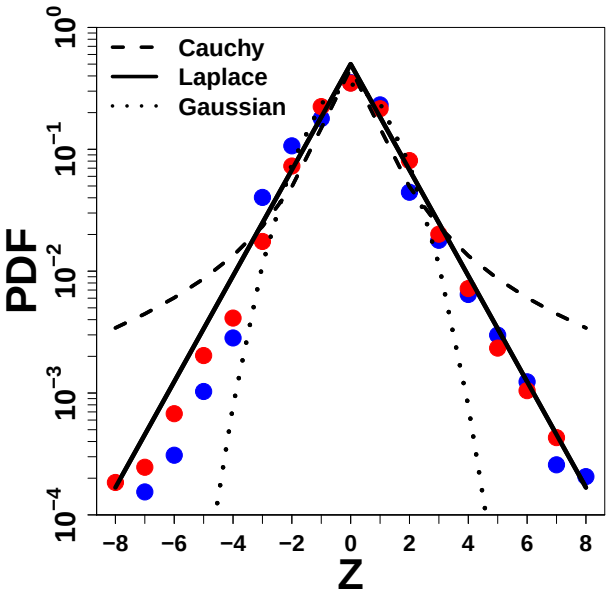

Figure 1.5: Distribution of fully normalized estimates $Z = \frac{X-m}{\sigma}$, before (blue) and after social influence (red) with all data from all experiments combined (close to 20000 estimates). The dots are the experimental values, and in black we compare the agreement of Gaussian (dotted line), Cauchy (dashed line) and Laplace (full line) distributions with the data. Without ambiguity, the Laplace distribution provides a much better approximation than the others.

where m is the center of the distribution (often called *location* parameter), σ is the width of the distribution (often called *scale* parameter) and n is the "tailedness" (often called *shape* parameter), which controls the thickness of the tails. The fatter the tails of a distribution, the higher the probability to find "outliers" (estimates very far from the center).

In 2010, Lobo and Yao studied various data sets of estimates and forecasts, and showed that the tailedness ranged from $n = 1$ (Laplace distribution) to $n = 1.6$, making the Gaussian distribution ($n = 2$) an actually quite poor approximation of most experimental distributions [129].

We now propose arguments that connect on the one hand Cauchy and Gaussian distributions, on the other hand Laplace and Gaussian distributions, to the amount of prior information individuals have about a question:

- Cauchy and Gaussian distributions belong to the so-called stable distributions family, defined as follows: $\{X_i\}$ being a set of estimates drawn from a symmetric probability distribution f characterized by its center m and width σ, we define the weighted average $X' = \sum_i p_i X_i$, with $\sum_i p_i = 1$. f is a stable distribution if X' has the *same* probability distribution f as the original X_i, up to the new width σ'. Indeed, the center m remains the same due to the condition $\sum_i p_i = 1$, but the width may decrease after averaging (law of large numbers), depending on the stable distribution f considered. Cauchy and Gaussian represent two extremes of the stable distribution family, Lévy distributions being intermediate cases: for the Cauchy distribution, the width σ remains *unchanged*, whereas the narrowing of σ is *maximum* for the Gaussian distribution (see also section 1.5.1). In the case of actual human estimates, the relevance of a certain distribution f can be related to the degree of *prior information* of the group. When individuals have no idea about the answer to a question, the weighted average of arbitrary answers cannot be statistically better ($\sigma' < \sigma$) or worse ($\sigma' > \sigma$) than the arbitrary answers themselves, leading to a Cauchy distribution for these estimates (the *only* distribution for which $\sigma' = \sigma$). However, when there is a good prior information, one expects that combining answers gives a better

statistical estimate ($\sigma' < \sigma$; Gaussian). Hence, when the quantity to estimate is very "easy" in the sense that it is closely related to general intuition (ages, dates, number of marbles in a jar... which was the kind of quantities generally asked in previous works), estimates should follow a Gaussian-like distribution, while when individuals have very little information about the answer, as in our experiments, estimates should be closer to be Cauchy-like distributed.

- The Laplace distribution maximises the entropy probability distribution under the constraint that the average absolute deviation from the median (σ; first moment) is fixed [130]. This means that for a certain value of σ, and if nothing else is known about the distribution, the most likely distribution to be observed is a Laplace distribution. There is a direct parallel with the Boltzmann distribution, which gives the probability distribution of energy states ϵ in a system of particles at fixed energy k_BT: $f(\epsilon, T) \propto e^{-\frac{\epsilon}{k_BT}}$, where k_B is the Boltzmann constant and T is the temperature of the system. The exponential distribution maximizes the entropy probability distribution for a fixed value of the average energy k_BT. The energy states ϵ play an equivalent role to the log-estimates X in our experiments, and the average energy k_BT to σ.

  Likewise, the Gaussian distribution maximises the entropy probability distribution under the constraints that the average (first moment) *and* the variance (second moment) are *fixed*. One understands that this additional constraint on the second moment is equivalent to a much larger amount of information in the system (the group).

  In other words, by asking a certain question to a certain group, the experimenter "fixes" some constraints, which depend on the amount of prior information detained by the individuals in the group about the question. If the group has very little prior information, one expects a minimal constraint, namely only a single constraint on the first moment: one would obtain a Laplace distribution. If the group has a much higher degree of prior information, one can expect an additional constraint on the second moment (no far outliers): one would obtain a Gaussian distribution.

In summary, the lower the demonstrability of the question, the closer the related distribution of estimates to a Laplace distribution, and the higher the demonstrability, the closer the distribution to a Gaussian distribution. The first argument above could appear a bit off-topic, since we showed that experimental distributions are better explained by Laplace distributions (GND) than by Cauchy distributions. Yet this argument is enlightening in the sense that it highlights the relationship between the distributions of estimates found and the degree of prior information of the corresponding groups. And it also remains theoretically valid for the Cauchy part at least: indeed, if one could imagine a question for which individuals in a group had no information *at all*, one would effectively find their estimates to be Cauchy distributed.

Remark: the Laplace distribution has another noticeable property: because of the optimality property of the median, it minimizes the absolute deviation from the median [131]. Since this quantity is a natural measure of diversity, it is tempting to say that the Laplace distribution minimizes diversity. Yet, as explained in introduction, decreasing diversity amounts to improve average individual accuracy, such that Laplace distributions in some sense maximize the average individual accuracy. However, we have not yet fully understood the deep meaning of this property, so we mention it as a remark but refrain ourselves from drawing hasty conclusions.

1.3.3 Laplace Distributions of Individual log-Estimates

The "full" normalization presented above provided very useful insights, in particular it made it very clear that the distributions of estimates for low demonstrability questions are Laplace-like. However, the collapse it forces onto the standard Laplace distribution forbids to quantify the

impact of social information on the process of estimation (indeed, both distributions before and after social influence collapse on the same standard Laplace distribution). To study this impact, we need to look at the distributions of the log-transformed estimates X (Figure 1.6):

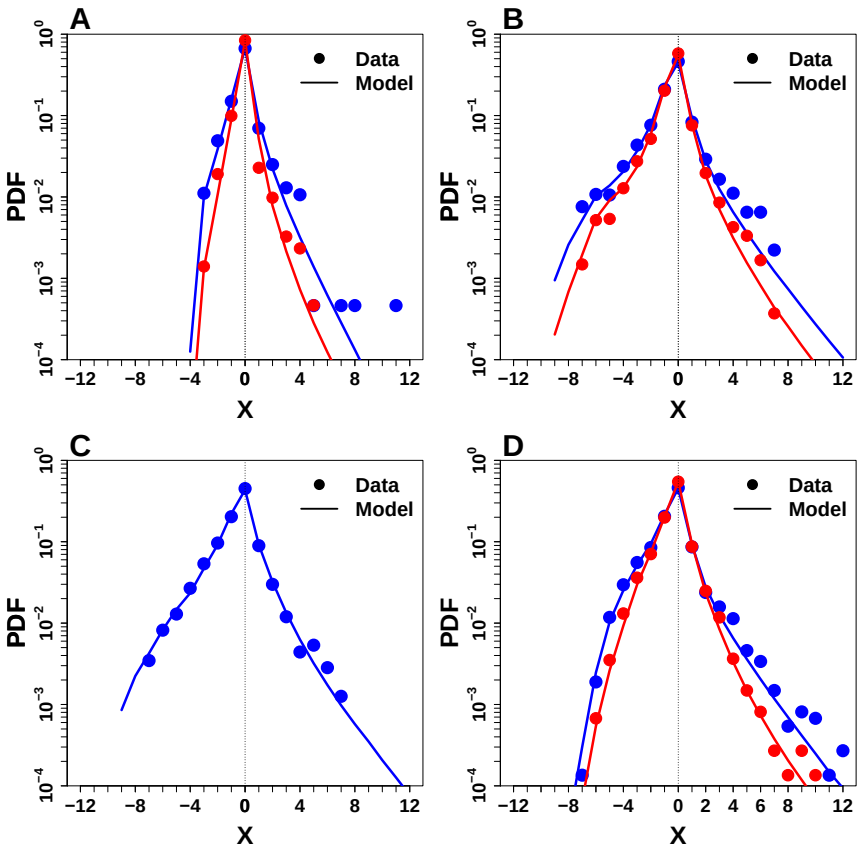

Figure 1.6: Distribution of log-transformed individual estimates $X = \log(\frac{E}{T})$, before (blue) and after social influence (red), in: A. the first experiment in Japan; B. the first experiment in France; C. the second experiment in Japan; D. the second experiment in France. The dots are experimental data while the lines are simulations from the model (detailed in section 1.5).

- The width of the distributions have decreased after social influence, confirming previous observations [58]. As explained in the introduction, this reduction in diversity also amounts to a collective improvement in accuracy, as estimates get overall closer to the truth (0 in log variables). Interestingly, this improvement occurred also in the second experiment in France, where subject were provided with incorrect information (Figure 1.6D; the impact of incorrect information will be studied in section 1.7). Here we focus on the "global" distribution (all values of ρ mixed), but this effect of information on the distribution's width actually depends on ρ. This aspect will be detailed in section 1.6.
- The distributions for the first experiments in France and Japan have very different widths. We hypothesized that this difference came from the difference in quantities to estimate. Indeed, Fig 1.7 shows that when the same questions were asked (second experiment in Japan), the distribution of estimates were very similar.
- There is a sharp decrease on the left side of each graph, due to the fact that for the chosen quantities to estimate, it doesn't make sense to provide an answer lower than one. Indeed, including this condition in our model (which amounts in log variables to $X > -\log(T)$; see section 1.5) reproduces the decay very closely. This effect also appears – although less clearly – in Figure 1.5 (distribution of Z for all four experiments), where the experimental points lie slightly below the theoretical line on the left parts of the graph.

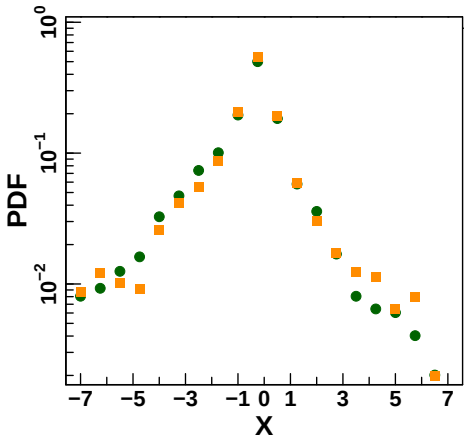

Figure 1.7: Probability distribution function (PDF) of log-transformed normalized estimates $X = \log(\frac{E}{T})$, in the first experiment in France (orange) and in the second experiment in Japan (green). The questions asked were the same, and the distributions are very similar, as predicted.

- Lower values than the truth are overweighted with respect to higher values, because of the human tendency to underestimate quantities.
- The distributions look different from usual Laplace distributions (straight lines in log-coordinates, as in Figure 1.5), due to the fact that linear combinations of Laplace random variables do not necessarily follow a Laplace distribution. As a consequence, even if for each quantity taken separately, estimates are Laplace distributed, the combination of estimates of several different quantities can result in a *curved* distribution as observed in the four panels (reminiscent of Cauchy distributions, which were our first hypothesis to replace Gaussian distributions).

1.4 Individual Sensitivities to Social Influence

The second quantity of interest in our study is the sensitivity to social influence: how do people use and react to social information? What are the mechanisms underlying the use of social information? Are there cultural differences? These are some of the questions we discuss in this section.

1.4.1 Distributions

After having received social information I_i, an individual i may reconsider her personal estimate E_{p_i}. One can always represent the new estimate E_i as the weighted average of the personal estimate E_{p_i} and the social information I_i. In the first experiment in Japan, the social information provided to the subject was the arithmetic mean of the τ previous answers (including that of the virtual "experts" providing the true answer $E_i = T$), because we assumed people would compute their new estimate E_i as the weighted arithmetic mean of their personal estimate E_{p_i} and the social information I_i. We therefore defined for this experiment the *sensitivity to social influence* S_i, by $E_i = (1 - S_i)\, E_{p_i} + S_i\, I_i$. $S_i = 0$ corresponds to subjects keeping their initial estimates, while $S_i = 1$ corresponds to subjects adopting the average estimate of their peers.

However, we realised later that humans tend to think in terms of orders of magnitude [121], and that the natural way for them to aggregate estimates is to use the median [122] or the geometric mean [60], which both tend to reduce the effect of outliers. Therefore, one should rather assume that humans compute the *geometric mean* of their initial estimate and the social

information, rather than the arithmetic one. Thus, in the two experiments in France, social information was defined as the geometric mean of the τ previous answers (including that of the virtual experts) $I_i = (\prod_{j=i-\tau}^{i-1} E_j)^{\frac{1}{\tau}}$, and the sensitivity to social influence S_i was defined by $E_i = {E_{p_i}}^{1-S_i} {I_i}^{S_i}$. $S_i = 0$ still corresponds to subjects keeping their initial estimates, and $S_i = 1$ to subjects adopting the average estimate of their peers.

In terms of log-transformed variables $X_i = \log(\frac{E_i}{T})$ (which are the natural variables to describe estimates), we obtain:

$$X_i = (1 - S_i)X_{p_i} + S_i M_i, \tag{1.5}$$

where the log-transformed social information is simply the arithmetic mean of the τ previous log-transformed estimates $M_i = \frac{1}{\tau}\sum_{j=i-\tau}^{i-1} X_j$, and thus:

$$S_i = \frac{X_i - X_{p_i}}{M_i - X_{p_i}} \tag{1.6}$$

Note that in this language, S_i is the barycenter coordinate of the final estimate in terms of the initial personal estimate and the social information.

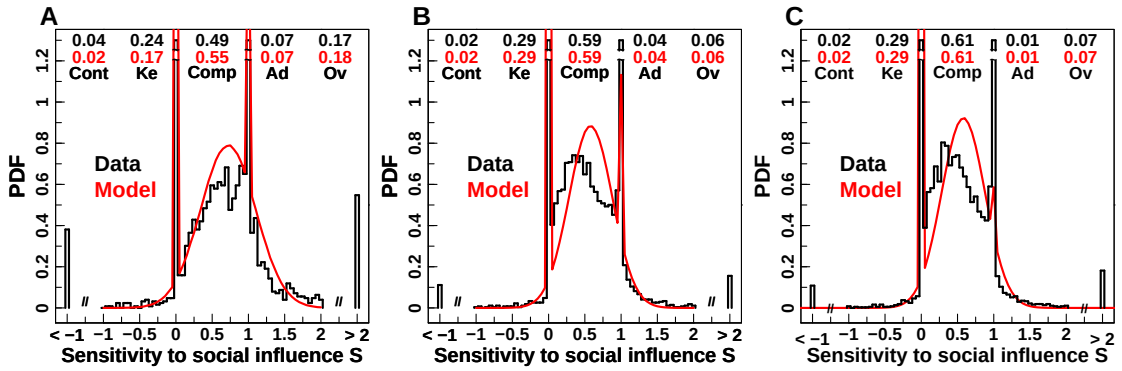

Figure 1.8: PDF of sensitivities to social influence S for: A. the first experiment in Japan; B. the first experiment in France; C. the second experiment in France (subjects only provided their personal estimates in the second experiment in Japan, so the sensitivity to social influence couldn't be measured). The numbers at the top of each panel are the probabilities for each category of behavior: contradict the social information (Cont; $S < 0$), keep one's opinion (Ke; $S = 0$), compromise (Comp; $0 < S < 1$), adopt the social information (Ad; $S = 1$), and overreact to it (Ov; $S > 1$). Experimental data are shown in black, and numerical simulations of the model are in red. The figure is limited to the interval $[-1, 2]$, and the values of S outside this range were grouped in the boxes $S < -1$ and $S > 2$. The box size is 0.05.

Figure 1.8 shows that the three experimental distributions of S have a similar shape: a bell-shaped part, that we roughly assimilate to a Gaussian, and two peaks at $S = 0$ and $S = 1$. The median value of S is about 0.35 in the experiments in France, in agreement with previous findings [60, 132, 119], suggesting that individuals tend to weigh personal information higher than social information [62, 118]. However, it is of about 0.6 in the first experiment in Japan, suggesting important cultural differences, and in particular that in Japan, the tendency is to give more weight to social than to personal information. It could be argued that the difference observed is mostly due to the difference in the questions asked in the experiments in France and Japan (higher demonstrability in Japan), which may have impacted the distribution. Yet, since other Western world works (previously cited) found similar median values to that measured in France (even though the quantities were different from ours, and in general highly demonstrable), we believe cultural differences are a better explanation.

Five types of behavioral responses can be identified: keeping one's opinion (peak at $S = 0$), adopting the group's opinion (peak at $S = 1$), making a compromise between the two ($0 <$

$S < 1$), overreacting to social information ($S > 1$), and contradicting it ($S < 0$). Quite surprisingly, responses that consist in overreacting and contradicting are generally neglected in the literature [123, 59, 132, 122], either considered as noise and simply not taken into account, or sometimes included into the peaks at $S = 0$ and $S = 1$, despite these behaviors being non-negligible (especially overreacting).

1.4.2 Consistent Differences in the Use of Social Information

In the first experiment in France, we analysed how consistent subjects' behavioral reactions would be: in each session, we split the subjects into three subgroups according to the way they modified their estimates on average in the first 24 questions. We first defined *confident* subjects as the quarter of the group minimizing $\langle |S_q| \rangle_q$, where q is the index of the questions (*i.e.* the subjects who were on average closest to $S = 0$) and the *followers* as the quarter of the group minimizing $\langle |1 - S_q| \rangle_q$ (*i.e.* closest to $S = 1$). The other half of the group was defined as the *"average"* subjects.

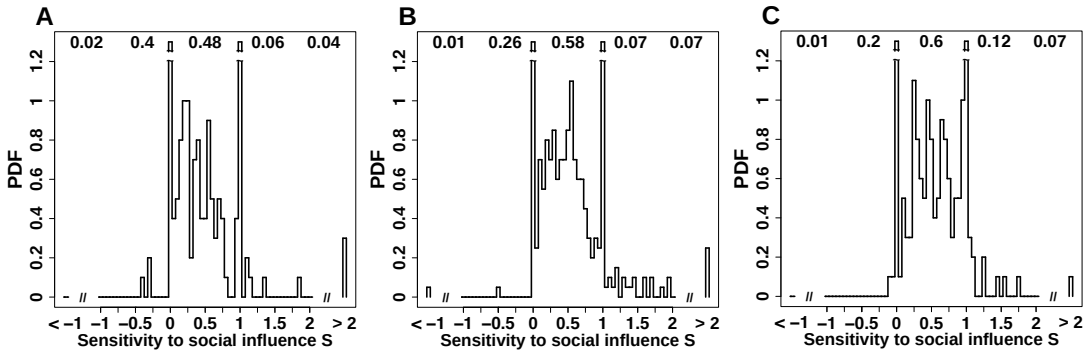

Figure 1.9: PDF of the sensitivity to social influence S, for questions 25 to 29 in the first experiment in France, in the following cases: A. confident subjects, defined as the quarter of each subgroup of 8 subjects minimizing $\langle |S_q| \rangle_q$, where q is the index for questions 1 to 24 (the subjects who were on average closest to $S = 0$); B. "average" subjects, defined as being neither confident nor follower; C. followers, defined as the quarter of the group minimizing $\langle |1 - S_q| \rangle_q$ (the ones who were on average closest to $S = 1$). In A, the peak at $S = 0$ is almost 7 times higher than the one at $S = 1$, whereas in B, it is less than 4 times, and in C, less than twice.

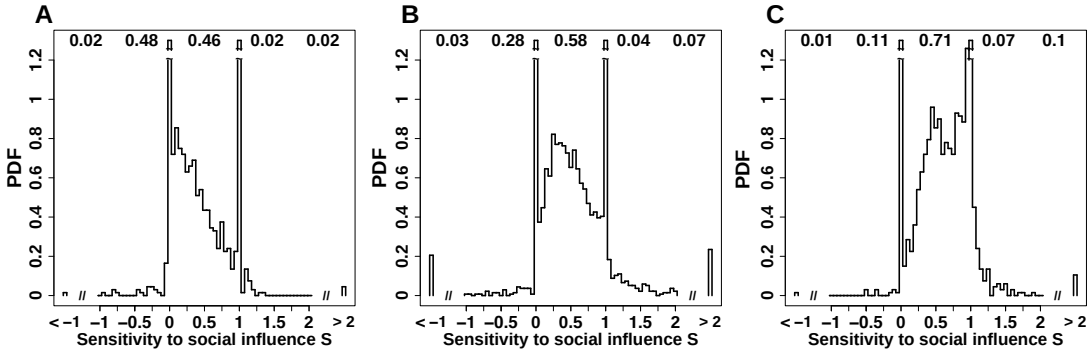

Figure 1.10: PDF of the sensitivity to social influence S, for individuals identified as confident (A), "average" (B) or followers (C). We reclassified individuals in these 3 categories, according to the same definition as given in Fig 1.9, but this time for questions 1 to 29 in the first experiment in France. The differences are more patent than in Fig 1.9, as expected.

Fig 1.9 shows the distributions of S for the three subgroups, computed from questions 25 to 29 (Fig 1.10 shows the same graphs, with the same categories recomputed for questions 1

to 29). The differences are striking, showing that subjects' behavioral reactions were highly consistent: for the group of confident subjects, the peak at $S = 0$ is about 7 times as high as the peak at $S = 1$, while for the group of followers, it is less than twice as high. Therefore, subjects who were characterized as confident from questions 1 to 24 remained highly confident, whereas subjects who were characterized as followers remained highly followers. Subjects who were characterized as "average" remained "average", in the sense that their distribution is very close to the global distribution (Fig 1.8 B). These results reflect robust differences in personality or general knowledge.

1.4.3 Distance between Personal Estimate and Social Information

Figure 1.11 shows that, on average, the sensitivity to social influence S depends on the distance $D = \log(\frac{E_p}{I}) = X_p - M$, between personal estimate (X_p) and social information (M).

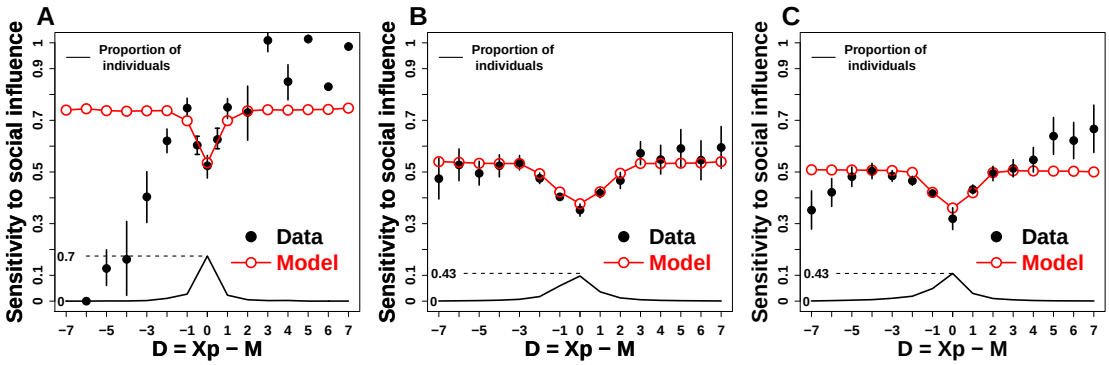

Figure 1.11: A. Mean sensitivity to social influence S against the distance between personal estimate and social information $D = X_p - M$, in: A. the first experiment in Japan; B. the first experiment in France; C. the second experiment in France. Full black circles correspond to experimental data, while red empty circles are simulations from the model. Beyond 3 orders of magnitude in the experiments in France are only about 14 % of data, and beyond 2 orders of magnitude in the experiment in Japan only 3.6 %.

Up to a threshold of t ($t \approx 1.5$ orders of magnitude in the first experiment in Japan, and $t \approx 2.5$ in the experiments in France), there is a linear cusp relationship between S and D: the farther away social information is from personal estimate, the higher the weight given to it (*i.e* S increases).

Fig 1.12 shows the origin of the correlation between $\langle S \rangle$ and D: as social information gets farther from personal estimate, the probability to keep one's opinion ($S = 0$) decreases, while the probability to compromise ($0 < S < 1$) increases. Interestingly, the probability to adopt does not change with D.

Notice that in the model we considered that beyond the threshold t, $\langle S \rangle$ becomes independent of D (plateau), disregarding thus the slight deviations from the plateau observed, mostly in Figure 1.11 A and C. Indeed, these deviations only concern a little fraction of the data, such that their consequences at the collective level are negligible. We therefore decided to neglect them in the model (see next section) as a first approximation, and it was indeed enough to reproduce all the main features observed in the data.

However, this effect is consistent in the three experiments, such that it deserves a few words: the asymmetry observed reflects a bias which consists in giving more weight – at equal (and far) distance from one's personal estimate – to social information (much) lower than one's personal estimate ($D = X_p - M \gg 0$) than to information (much) higher than personal estimate ($D = X_p - M \ll 0$). Notice that the tendency to underestimate quantities explains this phe-

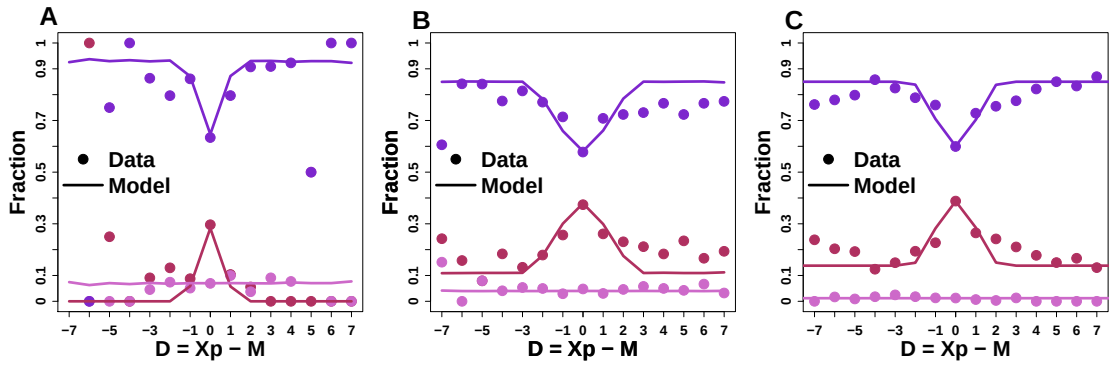

Figure 1.12: B. Fraction of subjects keeping (maroon), adopting (pink) and being in the Gaussian-like part of the distribution of S (mostly compromisers; purple) against the distance between personal estimate and social information $D = X_p - M$, in: A. the first experiment in Japan; B. the first experiment in France; C. the second experiment in France. Dots correspond to experimental data while full lines are simulations from the model.

nomenon: since people tend to underestimate quantities, lower values (than their own estimate) are perceived as more trustworthy (or plausible) than higher values.

As discussed in section 1.4.1, the differences observed between the experiments in France and Japan reinforce the cultural differences assumption, which is not the point of this thesis, but will be investigated in the future (see section 3.2.2 in the General Discussion).

In the next section, we present the model used in all the graphs shown in this chapter.

1.5 Model

Here we consider first a simple model, where the sensitivity to social information S is independent of $D = X_p - M$. Then we present the "full" model used all along this chapter. The interest of this preliminary model is that it can be solved analytically, shedding some light on the convergence of collective estimation processes.

1.5.1 Preliminary model

Let us consider first the case where "informers" are "experts", namely where they provide (the log-transform of) the true value $V = 0$. Equation 1.5 becomes:

$$\begin{aligned} X_i &= (1 - S_i)\, X_{p_i} + S_i\, M_i, \quad \text{with probability } 1 - \rho, \\ X_i &= 0, \quad \text{with probability } \rho \text{ (virtual experts)}, \end{aligned} \tag{1.7}$$

where X_{p_i} and X_i are respectively individual i's log-transformed personal and revised estimates, and the social information is $M_i = \frac{1}{\tau} \sum_{j=i-\tau}^{i-1} X_j$. "Experts" are assumed very confident so that their sensitivity to social influence is $S = 0$.

The distribution of personal estimates X_p is assumed stable[3], with stability parameter $a \in [1, 2]$, center (equal to the mean or the median for symmetric distributions considered here)

[3]In the actual "history" of our research, we initially assumed Gaussian distributions of estimates. Then, after the first experiment in Japan, we realised that the logarithm of estimates should be considered rather than estimates themselves, and that the distribution of log-estimates was closer to a Cauchy distribution than to a Gaussian distribution. Finally, it's only after the second experiment in France that we understood that Laplace distributions provided the best agreement with experimental data.

m_p and width σ_p. $a = 1$ and $a = 2$ correspond respectively to the Cauchy and Gaussian distributions, while intermediate values of a correspond to Lévy distributions. Consequently, the distributions of group's advice M and of estimates after social influence X are also stable (by definition of stability) with the same distribution but characterized by different center and width. We define m_i as the center and σ_i as the width of the distribution of estimates after social influence (excluding the virtual "experts"). From equation 1.7, for the i-th iteration/individual, and after averaging over configurations (*i.e.* an infinite number of experiments), m_i obeys the following recursion equations:

$$m_i = (1 - S)\ m_p + S\ (1 - \rho)\ m'_{i-1}, \tag{1.8}$$

where $m'_{i-1} = \frac{1}{\tau} \sum_{j=i-\tau}^{i-1} m_j$, and $S = \langle S_i \rangle$ is the mean sensitivity to social information. The term $(1 - \rho)\ m'_{i-1}$ is the mean social information provided by a fraction $(1 - \rho)$ of individuals and a fraction ρ of experts giving the estimate $V = 0$. For $i \to +\infty$, m_i and m'_{i-1} converge to the fixed point solution m_∞ of equation 1.8:

$$m_\infty = m_p\ \frac{1 - S}{1 - S(1 - \rho)}, \tag{1.9}$$

which is independent of τ. This analytical prediction of the simple model is plotted in Figure 1.13A, and is in fair agreement with the experimental data.

Similarly, the width of the distribution of the i-th individual estimates (excluding the virtual experts) satisfies the recursion relation (before averaging over configurations):

$$\sigma_i^a = (1 - S_i)^a\ \sigma_p^a + S_i^a\ (1 - \rho) \sigma'^a_i, \tag{1.10}$$

where $\sigma'^a_{i-1} = \frac{1}{\tau^a} \sum_{j=i-\tau}^{i-1} \sigma_j^a$. If personal estimates are Cauchy distributed ($a = 1$), then σ_i follows the same dynamics as m_i, and the asymptotic width is:

$$\sigma_\infty = \sigma_p\ \frac{1 - S}{1 - S(1 - \rho)}. \tag{1.11}$$

This analytical prediction of the simple model is plotted in Figure 1.13 B, showing a fair agreement with the experimental data.

Remarks:

- Note that this simple model (along with preliminary data from the first experiment in Japan) was exploited to design the experiment in France, in order to characterize the parameters for which significant results could be observed experimentally: number of subjects, percentage of virtual experts, number of sequential iterations and questions, values of τ...

- If the "informers" provide the value $V \neq 0$ instead of the true answer $V = 0$, the width of the distribution remains unaffected, but the asymptotic center of the estimate distribution becomes:

$$m_\infty = \frac{m_p(1 - S) + \rho\ S\ V}{1 - S(1 - \rho)}. \tag{1.12}$$

  We note that m_∞ can be driven to the exact result ($m_\infty = 0$) if the fraction ρ of virtual agents provides an *incorrect* information taking the optimal value:

$$V_0 = -\frac{m_p(1 - S)}{\rho\ S}. \tag{1.13}$$

Therefore, "informers" can lead the median estimate toward the true value (this will be discussed in detail in section 1.7), by providing an optimal piece of information $V = V_0$ strictly positive (since $m_p < 0$ due to the human bias to underestimate large quantities). This optimal V_0 becomes naturally larger as the density of virtual agents decreases (see Figure 1.18 for the confirmation of this effect in the full model).

- For $\tau = 1$, the full dynamics can be computed analytically as equation 1.8 reduces to $m_i = (1 - S)\, m_p + S\,(1 - \rho)\, m_{i-1}$. Defining $z = (1 - \rho)\, S < 1$, one obtains

$$m_i = z^i\,(m_0 - m_\infty) + m_\infty, \tag{1.14}$$

where m_0 is the initial condition. We hence find that m_i converges exponentially to its asymptotic value m_∞, a result which can be shown to remain true for any τ (m_i is then the sum of m_∞ and of τ exponentially decreasing terms).

1.5.2 Complete Model

Here we include the correlation between $\langle S \rangle$ and D, and we replace the stable distributions by Laplace distributions. The model simulates a sequence of 20 successive estimates performed by the agents ("informers" excluded). A typical run of the model consists of the following steps, for a given condition (ρ, α, τ) and a certain question q (we directly implement the log-transformed variables):

1. An agent's personal estimate X_{p_q} is drawn from the Laplace distribution $f(X_{p_q}, m_{p_q}, \sigma_{p_q})$, where m_{p_q} and σ_{p_q} are taken from the experimental distribution of estimates for question q. We impose that estimates are greater than one ($E_{p_q} > 1$) which amounts to $X_{p_q} > -\log(T_q)$ in log variables. This condition explains the fast decay on the left side of all distributions of log-estimates;

2. With probability ρ, an "informer" provides the value $V_q = \alpha_q \log(K_q)$ (the values of K_q are given in Table 1.1);

3. With probability $(1 - \rho)$, the agent receives as social information the average M of the τ previous estimates (estimates from "informers" included). $M_{0q} = V_q$ is chosen as initial condition;

4. The agent chooses its sensitivity to social influence S consistently with the results of Figure 1.8 and Figure 1.11.

   In particular, S is drawn in a Gaussian distribution of mean m_g with probability P_g, or takes the value $S = 0$ or $S = 1$ with probability P_0 and $P_1 = 1 - P_0 - P_\mathrm{g}$ respectively. P_0 and P_g have a linear cusp dependence with $D = X_p - M$, while P_1 is kept independent of D.

   For a given value of D, the average sensitivity to social influence is $\langle S \rangle = P_0 \times 0 + P_1 \times 1 + P_\mathrm{g} \times m_\mathrm{g} = a + b\,|D|$, where a and the slope b are extracted from Figure 1.11. P_g is hence given by $P_\mathrm{g} = (a + b\,|D| - P_1)/m_\mathrm{g}$.

   The threshold t is determined consistently by the condition $P_{\mathrm{g}_\mathrm{max}} = (a + b\,t - P_1)/m_\mathrm{g}$, where $P_{\mathrm{g}_\mathrm{max}}$ is the maximum value of P_g (plateau) beyond t in Figure 1.12. The values of all parameters are reported in Table 1.2;

5. The final estimate X_q after social influence is given by equation 1.5. The condition $X_q > -\log(T_q)$ is also imposed. One starts again from step 1 for the next agent.

	Jp 1	Fr 1 and Jp 2	Fr2
m_{g}	0.72	0.58	0.58
σ_{g}	0.38	0.32	0.3
P_1	0.07	0.04	0.012
$P_{\mathrm{g}_{\max}}$	$1 - P_1 = 0.93$	0.83	0.85
a	0.47	0.34	0.34
b	0.3	0.07	0.09

Table 1.2: Model parameters for the first experiment in Japan (first column; "Jp 1"), the first experiment in France (second column; "Fr 1") and the second experiment in Japan (second column; "Jp 2"), and the second experiment in France (third column; "Fr 2").

For all graphs, we ran $100,000$ simulations so that the model predictions' error bars are negligible (see Appendix 3.3 for their computation). Figure 1.8 shows that distributions of sensitivities to social influence S obtained from the model (red curve) are similar by construction to the experimental one. Also by construction of the model (step 4. above), the cusp dependence of the sensitivity to social influence on $D = X_p - M$ is well reproduced by the model (Figure 1.11; red curve and empty symbols).

We now address several non trivial predictions of the model.

1.6 Impact of Information Quantity on Group Performance

Consistently with our introductory discussion of the measurement methods of group performance, we propose the two following indicators, based on the median rather than on the mean (the estimators of the center and width of Laplace distributions being based on the median):

- *Collective performance*: $|\mathrm{median}(X_i)|$, which represents how close the center of the distribution is to 0 (the log-transform of the true value T)
- *Collective accuracy*: $\mathrm{median}(|X_i|)$, which is a measure of the collective proximity of individual estimates to the true value.

In this section we focus on the first experiments in France and Japan, which were devoted to the study of the impact of the quantity of information on group performance, given by the percentage of "experts" ρ present in the sequence of estimates. Indeed, their presence affects the value of the social information $M = \log(\frac{I}{T})$, and ultimately, of the final estimate X of the actual subjects.

1.6.1 Impact of Social Information on the Distributions' Center and Width

Figure 1.13 shows the collective performance (precisely defined above) and the width of the distributions of estimates, for the different values ρ and τ in the first experiment in France.

The collective performance is 0 when the distribution is centered on the true value, such that the closer it is to 0, the better. As expected, when $\rho = 0\,\%$, no significant improvement is observed in the collective performance. However, it is interesting to notice that the width decreases even for $\rho = 0\,\%$, namely when no external information was provided. It means subjects were able to use social information in a way that allowed a transfer of information from the most knowledgeable to the least knowledgeable individuals.

Then, as ρ increases, the center gets closer to the true value, and the width decreases accordingly, as also observed in the first experiment in Japan (see Figure 1.14). Notice that the

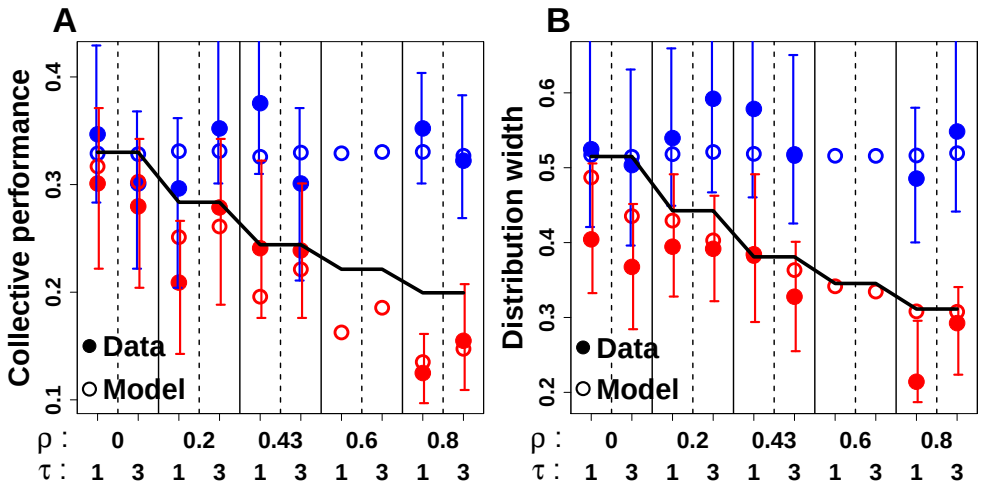

Figure 1.13: Collective performance (A; absolute value of the median of estimates) and width (B) of the distribution of estimates, for each couple of values (ρ, τ), before (blue) and after (red) social influence, in the first experiment in France. Both improve with ρ, as well as after social influence. Full circles correspond to experimental data, while empty circles represent the predictions of the full model. The full black lines are the predictions of the simple solvable model. For $\rho = 60\,\%$, only model predictions are available.

experimental error bars decrease after social influence, reflecting the decrease of the distribution width after social influence and the driving of people's opinion by the virtual "experts".

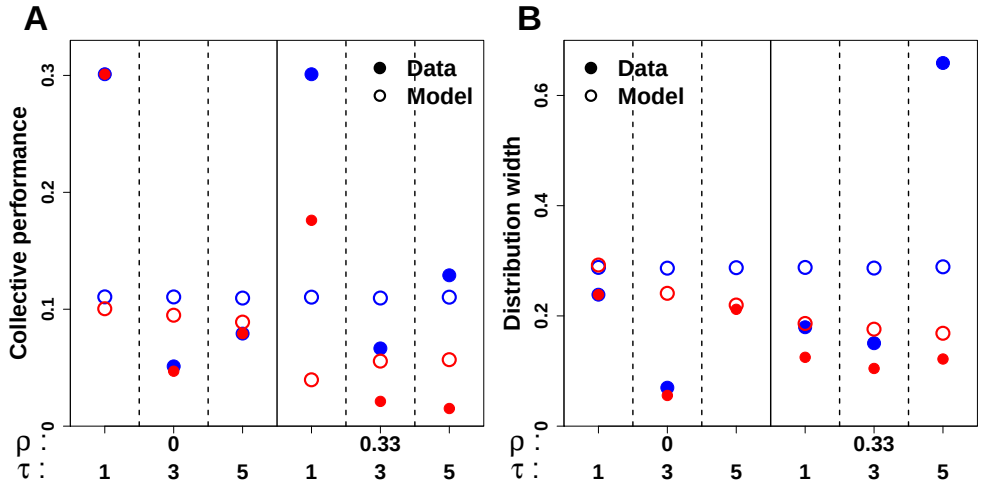

Figure 1.14: A. Collective performance (absolute value of the median of log-transformed estimates) in the first experiment performed in Japan, before (blue) and after (red) social influence, for each couple of values (ρ, τ). At $\rho = 0\%$ (no virtual experts), the performance doesn't change after social influence, as expected (the center of the distribution doesn't move). However, at $\rho = 33\%$, there is an improvement in collective performance after social influence for all values of τ. It is not clear whether there is an effect of τ, because the questions asked were different in all conditions (ρ, τ); B. Width of the distribution of estimates in the experiment performed in Japan, before (blue) and after (red) social influence, for each couple of values (ρ, τ). The distribution width has on average decreased after social influence, even at $\rho = 0\%$. The decreasing is nonetheless clearer at $\rho = 33\%$. Thus, the patterns are similar to those observed in the experiment performed in France (Fig 1.13). Note that because only two questions were asked in each condition, error bars could not be computed.

The collective performance and distribution width predicted by the model (empty circles) are in good agreement with those observed in the experiment. The very small effect of τ, only

reliably observed in the model, is explained in the remark below. The preliminary model (see description above) where we neglected the dependence of S on $D = X_p - M$, leads to fair predictions (full black lines in Figure 1.13), although it tends to underestimate the collective performance improvement and does not capture the reduction of the distribution width at $\rho = 0$.

Remark: our model predicts an effect of τ that is too small to be observed experimentally (much less than the experimental error bars). In our main model (like in the simpler solvable model developed above), the dynamics of estimates converges, although we cannot compute the convergence value analytically. The convergence value does not depend on τ, but the convergence speed (and hence the dynamics) does. It follows that the median estimate of 20 successive subjects, and hence the collective performance in Figure 1.13 A, depend on τ too.

The effect of τ in Figure 1.13 B can be understood intuitively: at $\rho = 0\%$, the larger τ, the more agents are likely to receive similar information, because of averaging effects. Therefore, subjects are also enticed to give more similar results, hence the smaller distribution width at $\tau = 3$ than at $\tau = 1$. This effect is the same when $\rho > 0\%$, but since experts are all providing the same information, the higher ρ, the more the experts' information takes over the agents' information, and hence the lesser the impact of τ.

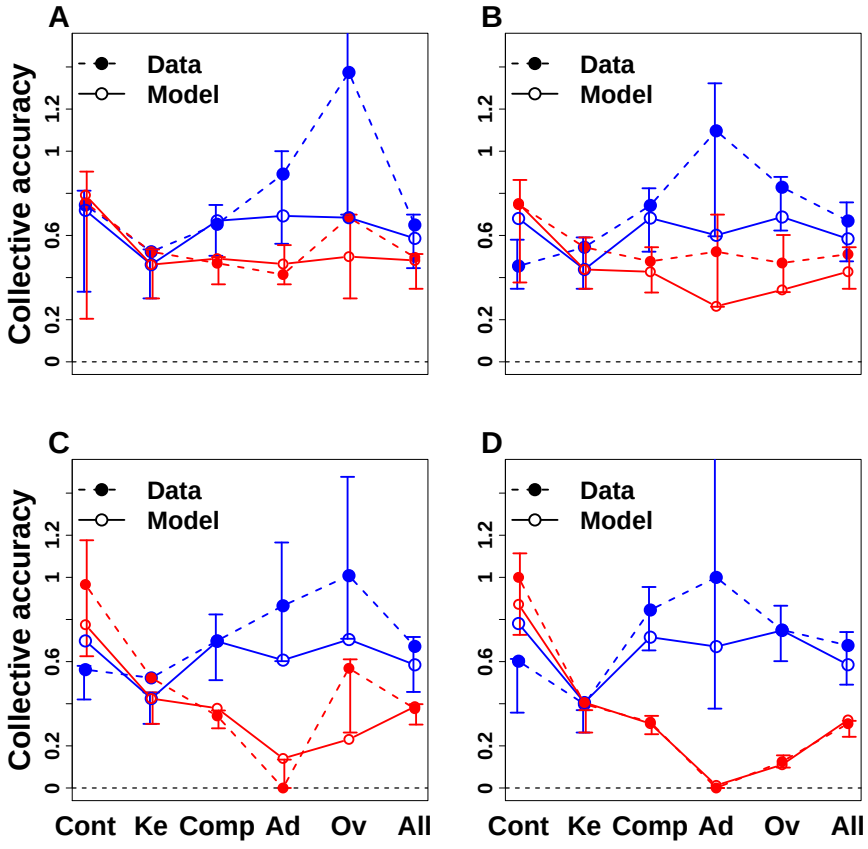

Figure 1.15: Collective accuracy (median distance of individual estimates to the truth) before (blue) and after (red) social influence, for the 5 behavioral categories identified in Figure 1.8 and for the whole group (All), for the first experiment in France: A. $\rho = 0\,\%$; B. $\rho = 20\,\%$; C. $\rho = 43\,\%$; D. $\rho = 80\,\%$. As ρ increases, the group improves its accuracy after social influence. Interestingly, the adopting behavior leads to the best improvement and accuracy after social influence for $\rho \geq 40\,\%$. Full circles correspond to experimental data, while empty circles represent the predictions of the model.

1.6.2 Impact of Social Information Use on Collective Accuracy

Figure 1.15 shows how collective accuracy (precisely defined above) improves after social influence as more reliable information is provided (namely as ρ increases), for each of the five categories of behavioural responses identified in Figure 1.8, as well as for the whole group.

At first thought, one would expect the collective accuracy *before social influence* (blue) to be independent of the response to social information. However, the data show (and the model predicts) that individuals who keep their opinion are significantly more accurate than others before social influence. The reason is quite subtle and a direct consequence of the cusp relationship between $\langle S \rangle$ and D (described in section 1.4.3): since the social information exchanged (geometric mean of previous estimates) must be statistically more and more accurate as τ increases (Wisdom of Crowds, see also Figure 1.3A and B), one expects individuals whose personal estimate is close to the social information (*i.e.* with a higher probability to keep; see Figure 1.11) to be on average more accurate than those whose personal estimate falls far from the social information (*i.e.* with a lower probability to keep; see Figure 1.11), at least when $\tau > 1$ ($\tau = 3$ here in half data). We expect this effect to be stronger for higher values of τ. A complementary explanation is that subjects who keep their opinion do so because they are quite confident in their answer, arguably because they have some knowledge about the quantity to estimate. Figure 5.10 in Appendix 3.3 shows the three ways correlation between sensitivity to social influence, confidence and accuracy.

After social influence (red), when no or little external information is provided ($\rho \leq 20\,\%$), subjects all perform about the same, which suggests that information flowed from the most accurate individuals to the least accurate ones. For $\rho > 20\,\%$, adopting leads to the best accuracy, while keeping and contradicting lead to the worst accuracy. Indeed, subjects who keep their opinion lose the opportunity to benefit from the information provided by the "experts", contrary to the subjects who use social information. Similar results have been found in the Japan experiment, as shown in Figure 1.16.

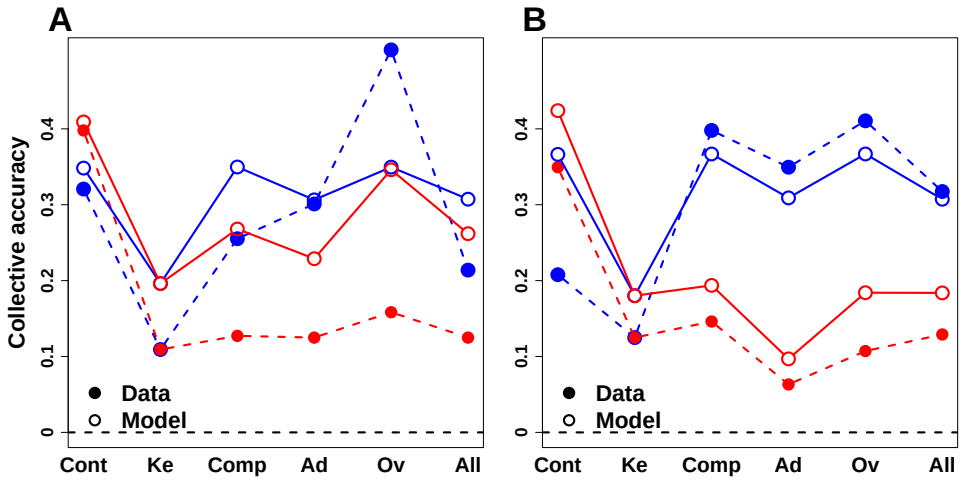

Figure 1.16: Collective accuracy (median distance of individual estimates to the truth) for the first experiment in Japan, before (blue) and after social influence (red), for the two values of ρ (percentage of experts): 0% (A) and 33% (B). Subjects who keep their opinion are the most accurate before social influence. In A, at $\rho = 0\%$, all behavioural categories lead approximately to the same accuracy after social influence, except contradicting, which leads to much worse accuracy. In B, at $\rho = 33\%$, adopting is the behaviour leading to the best accuracy after social influence. The group's accuracy (All) improves after social influence both without (A) and with (B) experts, and seemingly more markedly in the presence of "experts". However, we can't ascertain that this effect is effectively due to their presence, because the questions asked in A and B were different.

The collective accuracy for each behavioural category is again fairly well reproduced by the model (empty circles), the model becoming even more accurate as ρ increases. As explained in the case of Figure 1.13, the experimental error bars strongly decrease after social influence. In Appendix 3.3, Figures 4.2 and 4.3 show equivalent graphs for collective performance in France and Japan.

Notice that the information contained in Figure 1.15 is in fact 3-dimensional: percentage of experts ρ, behavioural category and collective accuracy (or performance; see Appendix 3.3). Figure 1.17 provides an alternative way to visualise the same information. In Appendix 3.3, Figure 4.4 shows the equivalent graph for collective performance in France.

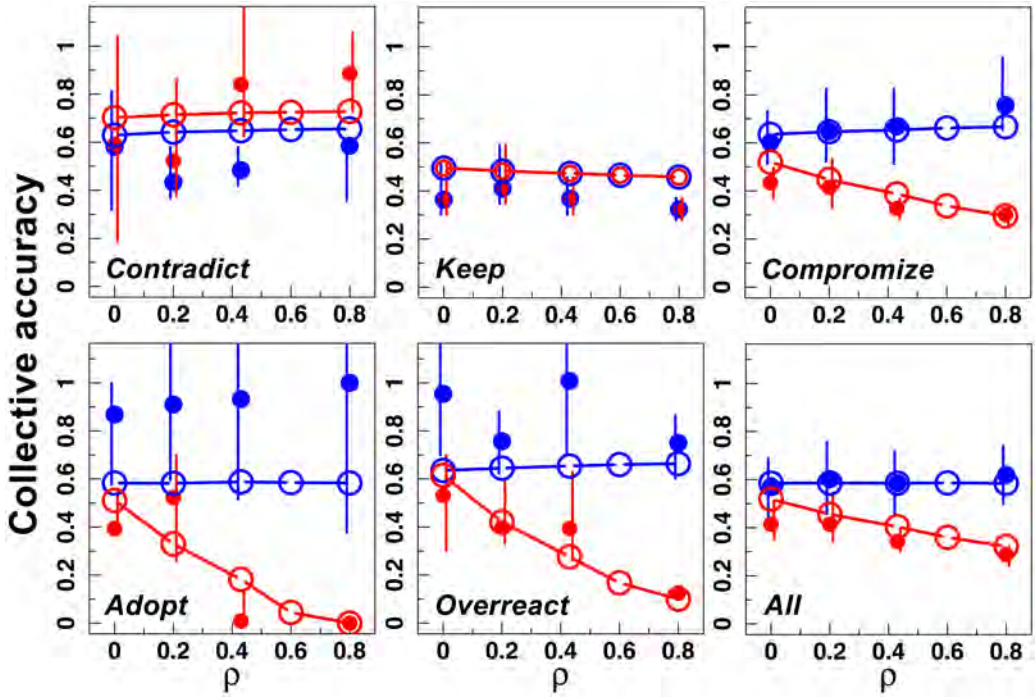

Figure 1.17: Collective accuracy (median distance to the truth of individual estimates) before (blue) and after (red) social influence against ρ, for the 5 behavioural categories identified in Figure 1.8 and for the whole group (All). Adopting leads to the sharpest improvement, and the best accuracy for $\rho \geq 40\,\%$. Full circles correspond to experimental data, while empty circles represent the predictions of the model (including for $\rho = 60\,\%$, a case not tested experimentally).

1.7 Impact of Information Quality on Group Performance

Prior to the second experiment run in France in 2017, we used the model to investigate the influence of the quality of information delivered to the group, *i.e.* the value V of the answer provided by the "informers", on collective performance and accuracy. We expected a deterioration of the collective performance and accuracy as V moves too far away from the truth 0, and as a greater amount of incorrect information is delivered to the group (namely as the quantity of "informers"ρ increases). The model predicted (see Figure 1.18) that the optimum collective performance and accuracy would be reached for a *strictly positive* value of V, whatever the density $\rho > 0$, as also predicted by our simple analytical model previously presented. Hence, incorrect information can be beneficial to group performance, contrary to the general intuition: providing overestimated values to the group can counterbalance the human cognitive bias to underestimate quantities [124].

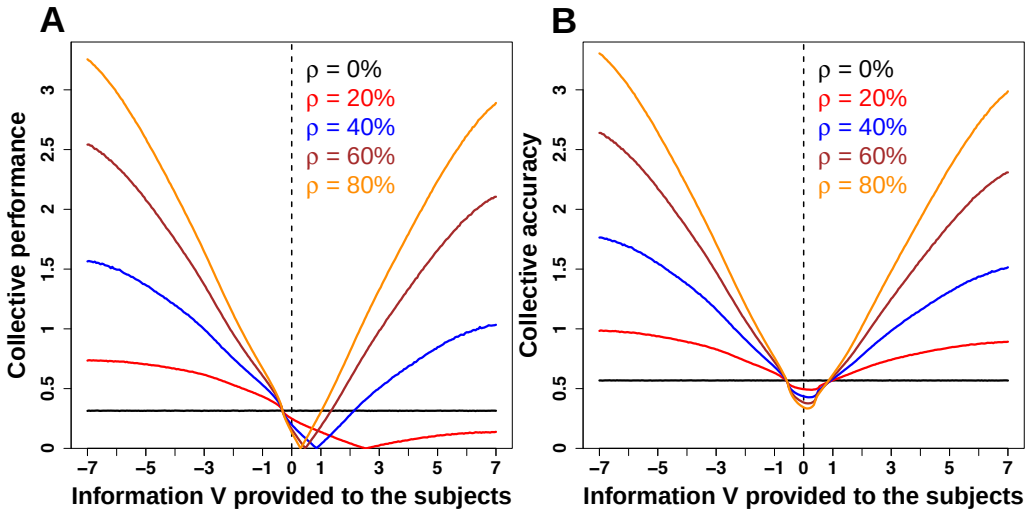

Figure 1.18: A. Collective performance against the value of the (log-transformed) information V provided by the "informers", for different values of ρ, as obtained using the full numerical model. For all values of ρ, we find that the collective performance reaches a minimum at a strictly positive value $V = V_0$ (decreasing as ρ increases), as predicted by the simple analytical model presented in subsection 1.5.1. Hence, providing incorrect information greater than the true value can be more beneficial (the collective performance closer to 0) to the group performance than providing the truth itself, by compensating for the human natural cognitive bias to underestimate quantities ($m_p < 0$). For $\rho = 20\%$, 40%, 60% and 80%, the optimum collective performance is respectively reached for $V_0 \approx 2.5$, $V_0 \approx 0.8$, $V_0 \approx 0.4$ and $V_0 \approx 0.25$; B. Collective accuracy, against the value of the (log-transformed) information V provided by the "informers", for different values of ρ, as predicted by the full model. The pattern is very similar to the one in A, the optimum being reached for $V > 0$, for any ρ. Note that the optimum collective accuracy improves (gets closer to 0) as ρ increases.

The second experiment in France – on which we focus in this section – aimed at verifying this prediction, as well as at probing in more depth the impact of incorrect information on individual and collective accuracy in human groups.

1.7.1 Incorrect Information can Help a Group Perform Better

As explained in the experimental protocol, the information provided by "informers" is quantified by α such that $T_I = T.K^{\alpha}$. In terms of log-variables, the information V provided by the "informers" is $V = \log(\frac{T_I}{T}) = \alpha.\log(K)$. α can be seen as a normalized value of V, this normalization being required by the fact that a piece of information T_I has a different impact according to the quantity to be estimated: for example, 3 orders of magnitude from the true value ($V = 3$) would seem obviously wrong for the age of death of a celebrity, while it would seem perfectly plausible for the number of stars in the universe, to anybody who is not an expert in cosmology. Since the dispersion of estimates σ_p correlates with a question's demonstrability (see Figure 1.3 and related explanations), $\log(K)$ can be thought of as an effective value $\sigma_{p_{\text{eff}}}$ of the dispersion of estimates for a certain question (the chosen values of $\log(K)$ are actually often very close to the corresponding values of σ_p). We will therefore from now on think of α as $\alpha = \frac{V}{\sigma_{p_{\text{eff}}}}$.

Since our quantifier of information quality is not V but α, we introduce a similar normalization in the definitions of collective performance and accuracy:

- Collective performance: $\langle|\text{Median}_i(\frac{X_{i,q}}{\sigma_{p_q}})|\rangle_q$;
- Collective accuracy: $\langle\text{Median}_i(|\frac{X_{i,q}}{\sigma_{p_q}}|)\rangle_q$.

Figure 1.19 shows that collective performance and accuracy after social influence depend on α, as expected.

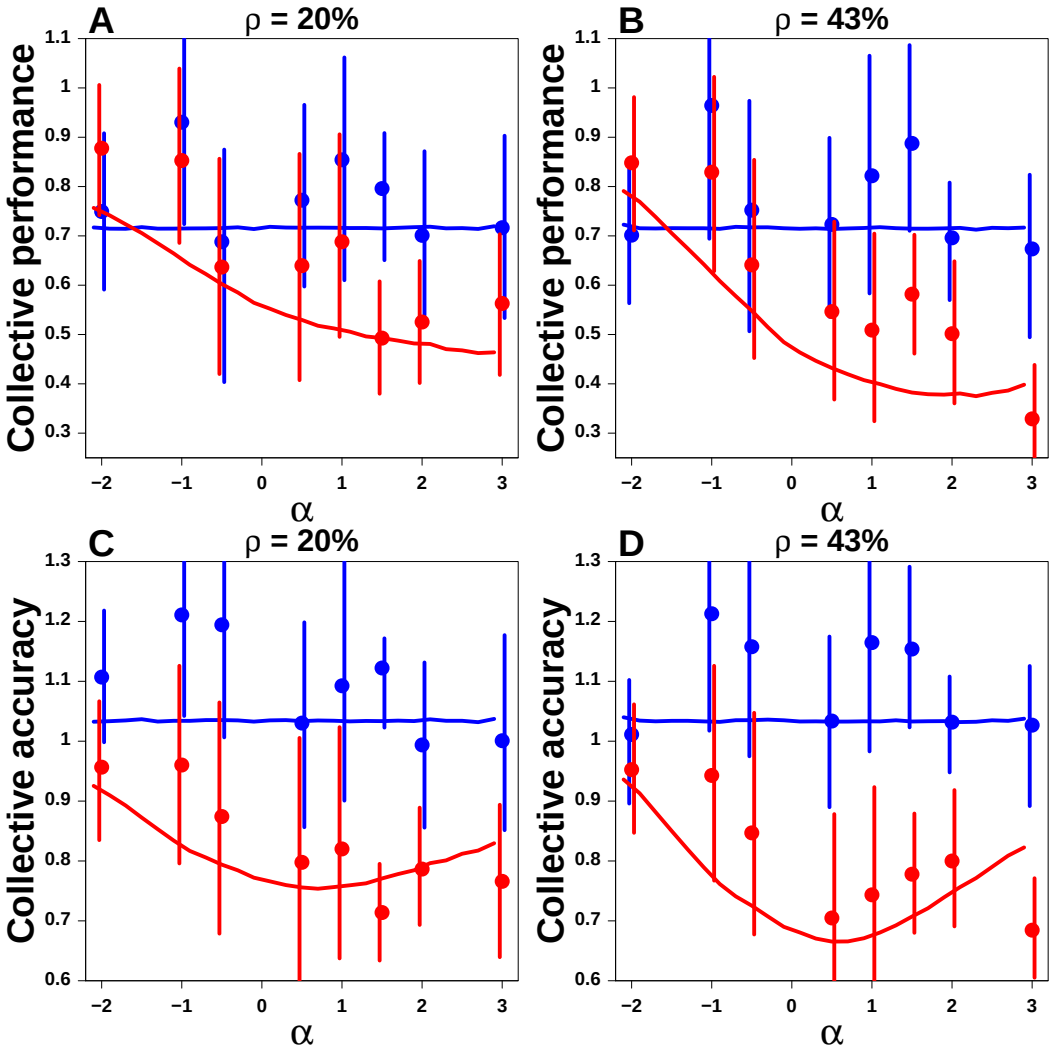

Figure 1.19: Collective performance (A and B) and accuracy (C and D) against $\alpha = \frac{V}{\sigma_{p_{\text{eff}}}}$*, before (blue) and after (red) social influence, for* $\rho = 20\,\%$ *(A and C) and* $\rho = 43\,\%$ *(B and D). Surprisingly, incorrect information can be beneficial to collective performance and accuracy, which optimums are reached for positive values of* α*. Dots are experimental data, and full lines are simulations from the model.*

More surprisingly, both improve (the dots and lines are closer to 0) after social influence (in red), *over almost the whole range of values of* α, which confirms the model's prediction that incorrect information can be beneficial to a group's performance. Only collective performance (Figure 1.19 A and B) for $\alpha < -1$ has decreased after social influence.

There are several interesting phenomena that need to be addressed step by step:

1. The crossing point of the blue (before social influence) and red (after social influence) lines on the left side of Figure 1.19A and B at $\alpha_{\text{left}} \approx -1.53$ is the value of the information provided α for which collective performance gets neither better nor worse. One would expect that α_{left} corresponds to the average absolute normalized bias of the group $\langle |\text{Median}_i(\frac{X_{p_{i,q}}}{\sigma_{p_q}})| \rangle_q \approx 0.73$ (collective performance before social influence; blue line). It is not trivial that providing the group with information farther away from the truth than the group's own bias can help the group improve its collective performance.

2. The optimum value α_{opt} of α for collective performance is strictly positive, as predicted. However, this optimum deviates from what one would expect from linear processes. Indeed, it would be natural to assume that the optimal value would be reached when α exactly compensates the normalized collective bias, namely $\alpha_{\text{opt}} \approx 0.73$. Yet, the data and the model suggest that the optimal value is much higher than 0.73 (and even much higher than -α_{left}) and depends on ρ, underlining the non-linear nature of the processes at play.

3. The impact of α is not symmetric with respect to its optimum. In particular, Figure 1.20 shows that the model predicts the incorrect information to be beneficial up to $\alpha_{\text{right}} \approx 13$ for $\rho = 20\,\%$ and $\alpha_{\text{right}} \approx 7.5$ for $\rho = 43\,\%$, which corresponds respectively to $0.46 \times 13 \approx 6$ and $0.46 \times 7.5 \approx 3.5$ orders of magnitude above the true value (we used that the average effective value of the distributions' widths over all questions q is $\langle \sigma_{p_{\text{eff}\,q}} \rangle_q \approx 0.46$, and then $\langle \sigma_{p_{\text{eff}\,q}} \rangle_q \times \alpha_{\text{right}} \approx V_{\text{right}}$).

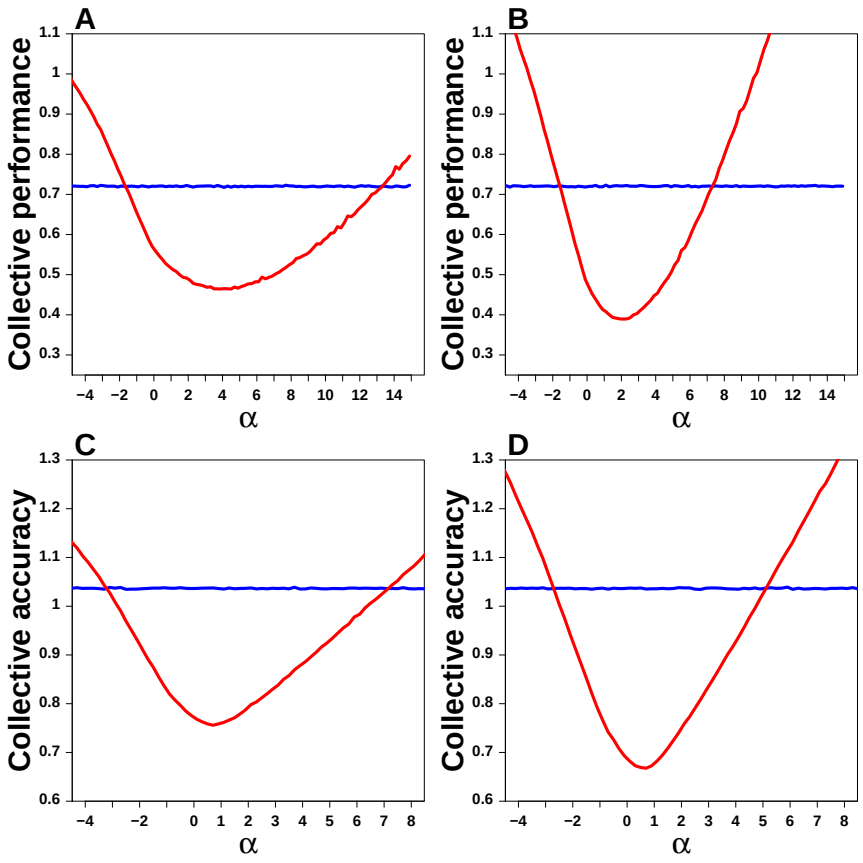

Figure 1.20: Collective performance (A and B) and accuracy (C and D) against $\alpha = \frac{V}{\sigma_{p_{\text{eff}}}}$, before (blue) and after (red) social influence, for $\rho = 20\,\%$ (A and C) and $\rho = 43\,\%$ (B and D). Only simulations from the model are shown, for a large range of values of α. Interestingly, even very incorrect information overestimating the truth can be beneficial to collective performance and accuracy.

Remarks:

- For the collective accuracy, the patterns and conclusions are similar, but the values are different.

- Our model predicts an effect of ρ on the collective performance and accuracy, but it is too weak to be observed experimentally. Note that it doesn't matter here, since the impact of ρ (quantity of information) was the focus of the previous section, not this one.

1.7.2 Incorrect Information and Sensitivity to Social Influence

We have seen in section 1.6 that collective performance and accuracy depend on subjects' sensitivity to social influence S (see Figure 1.15). Figures 1.21 and 1.22 compare collective performance and accuracy for three behavioral categories defined according to S:

- Subjects are *confident* when their S is in the lowest third;
- Subjects are *followers* when their S is in the highest third;
- Subjects are "*average*" otherwise.

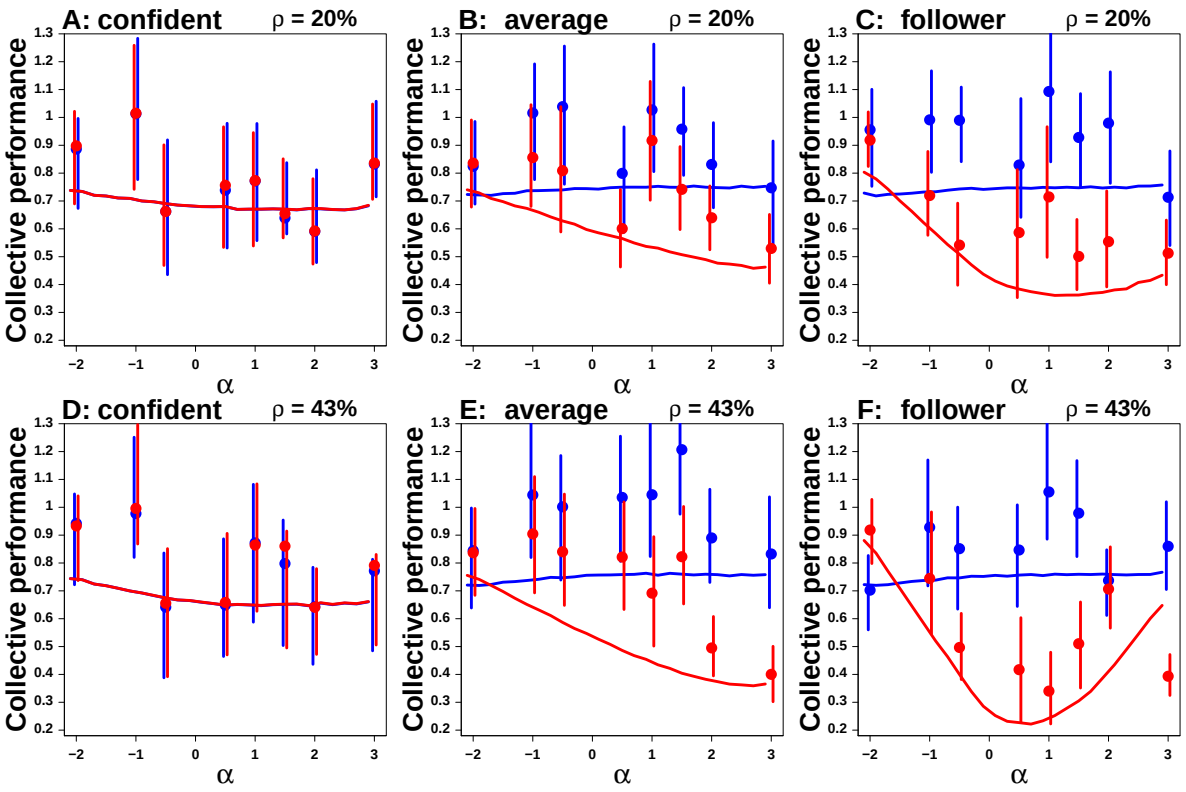

Figure 1.21: Collective performance against α, before (blue) and after (red) social influence, for $\rho = 20\,\%$ (A, B and C) and $\rho = 43\,\%$ (D, E and F). A. and D. Confident behavior; B. and E. "Average" behavior; C. and F. Follower behavior. Dots are experimental data, and full lines are simulations from the model.

The differences between the three behavioural categories are striking and provide useful insights into the use of social information in human groups:

- As expected, when subjects are confident (*i.e* tend to keep their opinion; Figures 1.21A, D and 1.22A, D), collective performance and accuracy don't depend on ρ, and are the same before and after social influence, because social information is disregarded. Likewise, α has little impact on collective performance and accuracy. There is a little effect observable in the simulations only because being confident doesn't mean exclusively keeping, but also sometimes contradicting ($S < 0$) or even slightly compromising (S slightly larger than 0).
- Collective accuracy was slightly (but non-negligibly) better for confident subjects than for the two other categories before social influence (blue lines). This effect is not observed in collective performance, and is thus related to the width of the distributions of estimates. We have seen before that the probability to keep one's opinion increases as personal estimate and social information get closer. Therefore, because of the redundancy

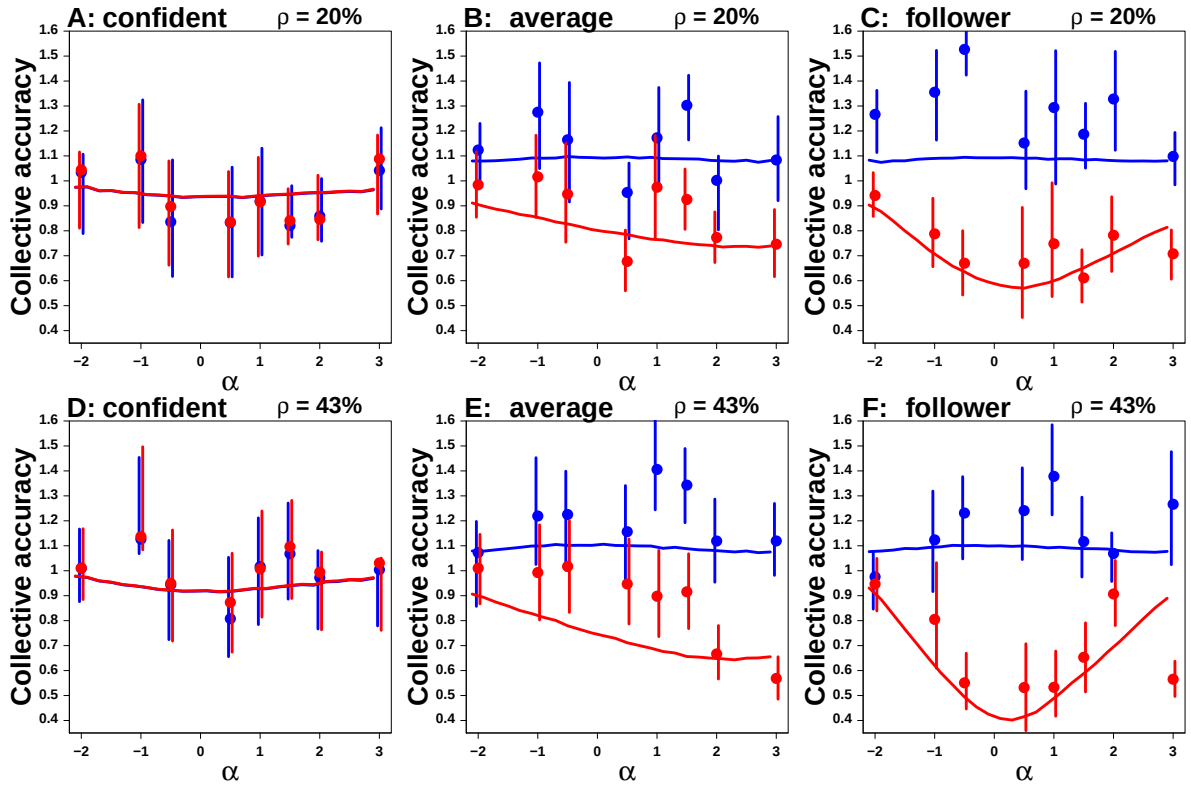

Figure 1.22: Collective accuracy against α, before (blue) and after (red) social influence, for $\rho = 20\,\%$ (A, B and B) and $\rho = 43\,\%$ (D, E and F). A. and D. Confident behavior; B. and E. Average behavior; C. and F. Follower behavior. Dots are experimental data, and full lines are simulations from the model.

in social information M provided ("informers" always provide the same value in a given treatment), there is also redundancy in the personal estimates of the confident behavior: those who keep their opinion often have personal estimates X_p close to M and hence to each other. Therefore the distribution of their personal estimates is narrower than for the other behaviours (but the center is the same) such that the collective accuracy is better.

- When subjects are "average" (Figures 1.21B, E and 1.22B, E), the impact of α is similar to its impact on the whole group, as expected.

- The most intriguing effect is that when subjects tend to follow the opinion of others (Figures 1.21C, F and 1.22C, F), they are able to reach better performance and accuracy than those who are "average" or confident, *even when some incorrect information is provided.* Following outperforms being "average" over a quite large range of values of α, and being confident over an even larger range.

1.8 Conclusion

Quantifying how social information affects individual estimations and opinions is a crucial step towards understanding and modeling the dynamics of collective choices or opinion formation [133]. Here, we have measured and modelled the impact of social information at individual and collective scales in *sequential* estimation tasks with *low demonstrability.*

More specifically, we rigorously controlled the information delivered to the subjects by means of "informers", virtual agents inserted into the sequence of estimates – unbeknownst to the subjects – and providing a value of our choice. The fraction ρ of "informers" defined the *quantity*

of information delivered, and the value provided its *quality*.

These virtual "informers" can be seen either as an external source of information accessible to individuals (*e.g.* Internet, social networks, medias...), or as a very cohesive (all having the same opinion V) and over-confident (all having $S = 0$) subgroup of the population, as can happen with "groupthink" [134].

The distributions of estimates in the four experiments we ran were closer to Laplace distributions than to Gaussian distributions as generally suggested in the literature. Gaussian and Laplace distributions are both members of the Generalized Normal Distributions family, and the proximity of experimental distributions of estimates to one or the other depends on the level of prior information held by individuals in a group about a quantity to estimate: when this prior information is poor (respectively high), the distribution of log-transformed estimates is close to a Laplace (respectively Gaussian) distribution.

We showed that after social influence, the center of a distribution of estimates gets closer to the true value (for $\rho > 0$) and its width decreases, which translates into an improvement in collective performance and accuracy, as defined in section 1.6. This improvement increases with ρ, namely with the quantity of information provided to individuals.

The distribution of sensitivities to social influence S is bell-shaped (contradict, compromise, overreact), with two additional peaks at $S = 0$ (keep) and $S = 1$ (adopt), which leads to the definition of robust social traits (confident, follower and "average" individuals), as confirmed by further observing the subjects inclined to follow these behaviours.

Our results revealed that when subjects have little prior information (low demonstrability questions), their sensitivity to social influence increases linearly with the difference between their personal estimate and the social information they receive, at variance with what was found in [62], for high demonstrability questions (see section 3.1.1 for a more detailed discussion about this). The mechanism behind this phenomenon is an increased tendency to compromise with social information as it is farther from one's own opinion.

As a direct consequence, individuals who tend to *keep* their opinion (often because it is close to social information) are on average more accurate than others before social influence. However, after social influence others manage to be as accurate when $\rho = 0\,\%$ (*i.e.* when no external information is provided), and even more accurate (except for the contradictors who are always the least accurate after social influence) when $\rho > 0\,\%$.

Interestingly, we found that individuals who tend to *adopt* the social information are almost perfectly accurate after social influence when $\rho > 20\,\%$, although they were the least accurate before. This suggests that crediting social information is the best strategy at least when the "environment" (the "informers" here) provides reliable information.

We built and calibrated a model of collective estimation that quantitatively reproduces the sharpening of the distribution of individual estimates and the improvement in collective performance and accuracy, as the amount of good information provided to the group increases.

We then studied the impact of incorrect information on collective performance and accuracy, for the whole group as well as for confident, "average" and follower behavioural subgroups (precisely defined in section 1.7.2). We defined a quantifier of information quality α as the (log-transformed) value V provided by the informers normalized by the effective dispersion $\sigma_{p_{\text{eff}}}$

of estimates for the question at hand (see the experimental protocol).

We found that providing incorrect information can help a group perform better than no information at all, and even better than providing the truth itself, by compensating for the human cognitive bias to underestimate quantities.

The pattern of this intriguing effect is not trivial and reveals non linear effects. In particular, the impact of α is not symmetrical: collective performance can be improved by delivering incorrect information overestimating the truth up to several orders of magnitude, whereas it degrades much faster if the information delivered underestimates the truth.

Finally, we showed that following social information leads to the best performance and accuracy after social influence, over quite a large range of values of α, suggesting that trusting others is the best strategy not only when the information available (delivered by "informers") is perfectly accurate ($V = 0$), but also in uncertain environments where the information available is misleading (up to a certain point).

This conclusion is reinforced by the fact that confident people are generally better than others before social influence, but perform worst after social influence, even when the information provided is very far from the truth – the precise cognitive and behavioral mechanisms underlying these fascinating results are still under investigation.

Overall, we found that individuals, even when they have very little prior information about a quantity to estimate, are able to use information coming from their peers or from the environment to improve their individual and collective accuracy, as long as this information is not too highly misleading.

Ultimately, getting a better understanding of these influential processes opens new perspectives to develop information systems aimed at enhancing cooperation and collaboration in human groups, thus helping crowds become smarter [135, 136].

Chapter 2

Impact of Filtered Information on Human Phase Separation Processes

2.1 Introduction

In many animal societies, individuals are able to collectively self-organize to accomplish very complex tasks, such as foraging [137], hunting [138], building nests [139], avoiding predators [140, 90, 141], etc. These complex patterns occur without external control and emerge from local interactions among individuals [142, 70]. Because they have limited computational abilities and information about their surroundings, behavioral responses in many animals are often triggered locally by information coming from close neighbors. Thus, local alignment, attraction and repulsion among fish trigger collective swarming, schooling or milling behaviors [87, 88], flocks of birds can change direction very quickly and cohesively by following the movements of their closest neighbors [89, 90], and simple mimetic rules lead to herding behavior in sheep [91, 92].

Similar phenomena have been observed in human societies, such as in car traffic [143], urban organization [144] or pedestrians motion [102]. Collective patterns such as lane or trail formation [102, 145, 146, 147], or the dynamics of crowd disasters [95, 148, 149, 150] emerge from local interactions among individuals [66]. Pedestrians are thus a very fertile soil for the study of collective behavior, and a lot remains to be understood before being able to, for example, drastically reduce the extent and probability of disasters in crowd panics. Partial solutions have been proposed, such as improving evacuation systems or designing infrastructures to facilitate the flow at bottlenecks or crossroads. We believe that a complementary approach based on the interaction of pedestrians with *information-filtering systems* may be fruitful. However, how such systems could interact with and steer collective pedestrian behavior is still largely unknown.

To tackle this issue, we designed specific tasks in which pedestrians had to rely on such an information-filtering system: pedestrians in groups[1] of 22 individuals (confined in a circular arena), were randomly assigned a "color" (sub-group), but knew neither their own color nor that of other group members (no visual cue). The task was to segregate in clusters of the same color (*phase separation*).

To help pedestrians complete their tasks (because they had no visual cue), we designed a system mimicking a *sensory device* (*e.g.* the retina), able to condensate the full information accessible (colors and positions of all pedestrians in the arena) into a bit of information: tags attached to subjects' shoulders altogether transmitted the positions and colors of all pedestrians

[1] To avoid any confusion between terms, we use the word "group" for the whole set of participants involved in experiments (1, 2 or 22), "sub-groups" for sets of 11 individuals of the same color in an experiment, and "clusters" for groups of individuals (of any size) of the same color gathered at any time in the segregation process.

in real time to a central server, and delivered in return an acoustic signal (a "beep") if specific conditions were satisfied:

- *majoritarian* rule: a subject's tag would beep if the majority of her k (k =1, 3, 5, 7, 9, 11, 13) nearest neighbors were of a different color from her.

- *exclusive* rule: a subject's tag would beep if *at least one* of her k (k =1, 2, 3, 4) nearest neighbors were of a different color from her.

- *shifted* rule: a subject's tag would beep if the majority of her k^{th}, $(k+1)^{\text{th}}$ and $(k+2)^{\text{th}}$ (k =1, 2, 3, 4) nearest neighbors were of a different color from her.

Subjects were not aware of the precise rules, and were simply told that they would "beep" whenever their "environment" would be of a different color. By varying k in the *majoritarian* and *exclusive* conditions, we controlled the amount of neighbors considered in the computation of the beep, and hence the amount of information processed. By increasing k in the *shifted* and *exclusive* conditions, we manipulated the task "complexity" (notice that we don't change the amount of information processed in the *shifted* condition).

We defined group performance as the segregation time (shorter time is better) and, by analogy with phase separation phenomena, the number of clusters at final time (less clusters is better). We looked at the impact of k on these two measures in all conditions, and found a U-shaped dependence of the segregation time on k, while the average number of clusters at final time decreased with k, and reached a plateau at values of k that vary with group size.

To better understand the mechanisms underlying these results, we built and calibrated a model of pedestrian motion based on the *social forces* approach [99]. This approach has been very successful in first describing qualitatively collective phenomena such as those above mentioned [95, 148, 150], but the interaction forces were generally assumed and calibrated by fitting experimental data [151, 152]. Later, efforts have been made to *extract* the form of the interactions from the data and thus improve the precision of the social force models [103]. We follow-up in this direction and propose to go a step further in the quality of the description of the interactions, by combining recent methods that have proven successful in fish [153, 107].

The general approach consists in collecting motion data for single individuals (spontaneous motion and interactions with the physical environment), then for two individuals only (interactions between individuals) and finally for groups. From the data with one and two individuals, one can define a model, which predictions for collective motion are then compared to the data obtained from the whole group [153]. In our case, groups of 1, 2 or 22 pedestrians were asked to walk as randomly as possible in the circular arenas (groups of 22 individuals were asked to do so for 45 seconds before the beeps would start). Then, to build the model, we used a procedure (detailed in Appendix 3.3) allowing us to extract the actual functional form of the interactions directly from the data [107].

Although the investigation of the interaction forces between pedestrians is not complete yet, the current version of our model is already able to reproduce very closely many different quantities used to characterize collective motion: speed autocorrelation, (bounded) diffusion process and distributions of speeds, positions, closest neighbors and angles to the wall. The same model of pedestrian motion successfully reproduced collective patterns observed in the segregation phase, and allowed us to make predictions for segregation processes at different groups sizes.

2.2 Experimental design

Two series of experiments were conducted in September 2015 and June 2016 at Paul Sabatier University in Toulouse, France. A total of 115 participants in 2015 and 209 participants in 2016 took part in the study. They were naive to the purpose of our experiments and gave written and informed consent to the experimental procedure. Each participant was paid 10 Euros per hour, whatever the number of sessions he or she participated in.

The participants had to walk in closed circular arenas, either alone or in groups of 2 or 22, according to predefined rules that were given at the beginning of each series of experiments. Single or pairs of individuals had to walk randomly for about 3 min, while groups of 22 individuals started with a random walk for about 45 s before starting a "segregation phase" (explained in detail below).

The aims and procedures of the experiments have been approved by the the Ethical Evaluation Committee of INSERM (IRB00003888, decision n° 15-243).

2.2.1 Experimental setup

The experiments were performed in a large hall where circular arenas of radii 1.78 m (≈ 10 m^2; yellow tape), 2.52 m (≈ 20 m^2; red tape) and 3.56 m (≈ 40 m^2; blue tape) were marked on the ground (see Fig. 2.1A). Only the largest arena was used for groups of 22 individuals. Individual trajectories were tracked using a real time location system (RTLS) developed by Ubisense™, based on Ultra-Wide Band (UWB) signals triangulation. When compared to camera-based tracking of individuals, the fundamental feature of UWB-based localization is that the tracking of each individual in a group is 100% accurate: each individual is assigned a uniquely defined pair of tags that are unambiguously associated with the spatial location of the individual.

The tracking system consisted of a large set of tags, attached to the participant's shoulders, that emitted Ultra-Wide Band wave trains, and sensors precisely placed in order to cover the experimental arena, that received and processed signals from tags. In our experiments, the tracking system included 6 to 8 sensors uniformly distributed around the arena, fixed 2-3 m from the border of the arena and 4 m from the ground. Ubisense sensors, as depicted in Fig. 2.1, are UWB signal receivers linked together and with the server by high-speed low-latency Ethernet connections. These devices are rectangular boxes of size $21.5 \times 15 \times 9$ cm^3 and weight 1 kg. All these sensors were wired-connected through a router to a server that actually performed the localization of tags and was able to send back some information to the tags.

Ubisense tags are miniaturized circuits powered with batteries that operate both in the 6 to 8 GHz frequency band (UWB signals used for localization) and in the 2.4 GHz band (signals used for data exchange and synchronization). Each participant was wearing two tags, one on each shoulder ($8.3 \times 4.2 \times 1.1$ cm^3 and 32 g for the left tag, $3.8 \times 3.9 \times 1.65$ cm^3 and 25 g for the right tag), attached by clips. These tags can emit an acoustic signal (a beep) triggered by the central server that has a global view of all participants' locations and colors.

Sensors were tightly synchronized together to measure the positions of tags. Each sensor captured the electromagnetic wave signals (pulses) emitted by the tags and measured in real time the angle from each tag to the sensor with an angular accuracy of 1.5 degrees. Additionally, pairs of sensors computed the time difference of arrival for a given tag pulse train. The server received by a wired Ethernet connection all available information (angles and time differences of arrival for each tag emission) and was able to derive the positions of the tags.

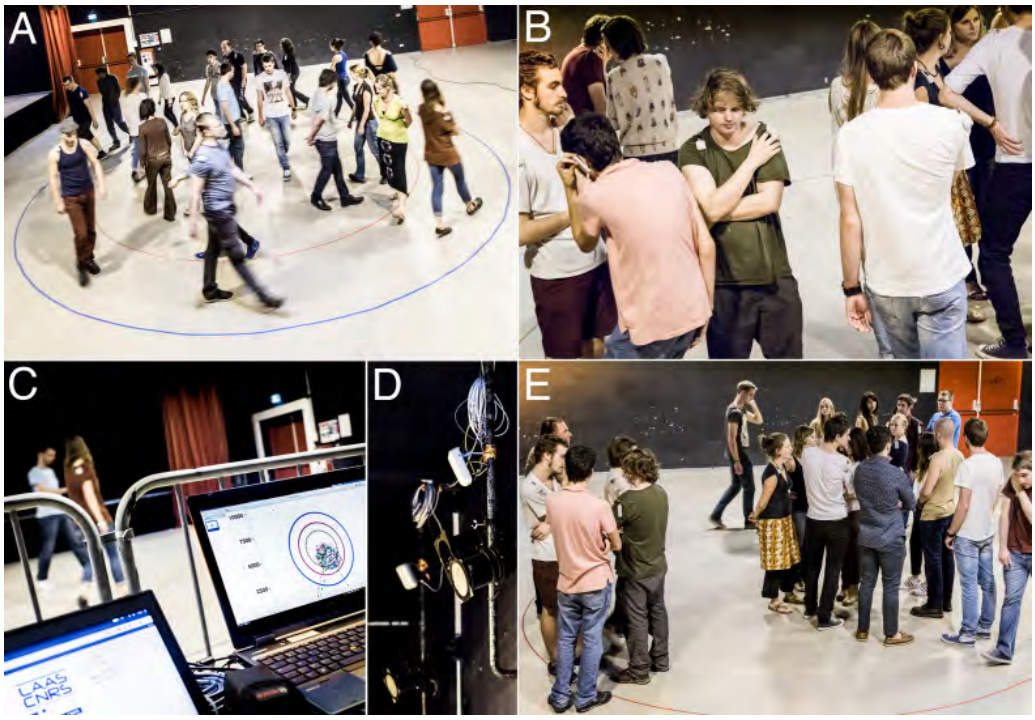

Figure 2.1: Experiments with 22 individuals and tracking system. The experiments started with the random assignment of colors to the 22 participants (11 blue and 11 red) and the random walk phase (A), where participants walked in the larger circular arena (delimited with a blue tape on the ground). In the segregation phase (B, E), the tags attached to the participants' shoulders emitted acoustic signals according to the conditions used to specify the environment of a participant. In the majoritarian case shown here, a tag emitted an acoustic signal to a participant if the color of the majority of his k nearest neighbors was different from his own color. The positions of participants were recorded in real time with a Ubisense tracking system (C, D) based on the triangulation of sensors (D) located around the arena.

In our experiments, the information collected by the sensors was used to evaluate in real time the environmental conditions of each subject by processing the relative position and color of all the subjects, and to react according to defined experimental conditions. These conditions described hereafter determine the state of a tag, that is, whether it should emit the acoustic signal or not. The information about the acoustic signal (beeping or silent) was sent back to each individual with a frequency of 1 Hz so that individuals could react to this information in real time with a short delay of less than one second. The acoustic signal emitted by a tag is of short duration and low intensity and could be unmistakably perceived by the participant who carries that tag. The system was set-up in order to perform the real-time double tracking (two tags per individual) of 22 individuals, whose instantaneous positions were determined with an error of less than 30 cm and recorded with a frequency of 2 Hz for each tag. In addition to the tracking system, the experiments were recorded with two cameras: a 360-degree camera fixed to the ceiling of the experimental room, and a camera located at four meters from the floor providing a lateral view.

2.2.2 Experimental protocol

Our aim was to build an interface that controlled the amount of information processed by the information-filtering system before a beep was provided locally to the 22 individuals performing the segregation task, in order to determine under which conditions this segregation process could be optimized in time or according to criteria such as segregation quality (phase separation).

We tested a total of 6 experimental conditions. Hereafter we describe these conditions and the instructions provided to the participants.

Environmental conditions tested in the experiments

Segregation experiments consisted of 381 sessions summarized in Table 2.1 and described in detail hereafter. Three kinds of segregation tasks were considered, depending on the criterion used to define the environment of a participant that triggered its tags to emit the acoustic signal.

In the first condition we used a *majoritarian* criterion: the tag carried by a participant would emit an acoustic signal if the color of the majority of his k nearest neighbors was different from his own color (see Fig. 2.2A) We used odd values for k (k = 1, 3, 5, 7, 9, 11 and 13), in order to have a proper definition of the majority.

Majoritarian k :	1	3	5	7	9	11	13	Total
Classic segregation	27 (22)	26 (20)	27 (22)	27 (21)	25 (19)	24 (19)	28 (0)	184 (123)
2 Clusters	17	17	17	17	17	17	17	119
Exclusive		$k = 2$		$k = 3$		$k = 4$		
Classic segregation		9		10		1		20
2 Clusters		10		9		3		22
Shifted		234		345		456		
Classic segregation		12		12		12		36

Table 2.1: The 381 sessions of segregation, arranged according to the beeping criterion. Majoritarian: the tag beeps if the majority of the k nearest neighbors are of the other color. Exclusive: the tag beeps as soon as one of the k nearest neighbors is of the other color. Shifted: the tag beeps if at least two of the three k-th, $(k+1)$-th and $(k+2)$-th neighbors are of the other color. 2 Clusters refers to the case where participants were explicitly asked to form two clusters at the end of the session. Numbers between parentheses denote sessions carried out in September 2015.

In the second condition we used an *exclusive* criterion: the tag carried by a participant would beep if *at least one* of his k nearest neighbors was of the other color (see Fig. 2.2B). This criterion is more difficult to satisfy, and we only used k = 2, 3 and 4; for $k \geq 5$ the numerical simulations of our model predicted excessively long segregation times (tens of minutes). Note that the case $k = 1$ is equivalent to using $k = 1$ in the previous majoritarian case.

In the third condition, we used the majoritarian criterion in a *shifted* environment composed of the k-th, $(k+1)$-th and $(k+2)$-th nearest neighbors, with k = 2, 3 and 4 (see Fig. 2.2C). For k = 2, the tag carried by a participant would emit an acoustic signal if the color of the majority (*i.e.* at least two) of the 2$^{\text{nd}}$, 3$^{\text{rd}}$ and 4$^{\text{th}}$ nearest neighbors was different from his own color. These cases are denoted '234', '345', '456' in Table 2.1 for k = 2, 3 and 4 respectively (in reference to the rank of the neighbors considered). Note that the case $k = 1$ is equivalent to $k = 3$ in the majoritarian case. For the majoritarian and the exclusive cases, we performed a second series of experiments where participants were explicitly asked to form "two clearly visually distinguishable clusters[2]".

Fig. 2.2 shows that the tag of a focal individual can beep or not depending on the beeping criterion and the value of k. In the examples of the majoritarian criterion (Panel A), the tag is silent for $k = 1$ and for the first example of $k = 5$, but beeps for $k = 3$ and 7 and for the second

[2]Notice that the word "cluster" doesn't have a proper translation in French, such that we actually asked participants to segregate in two "groups", without any possible confusion at their level.

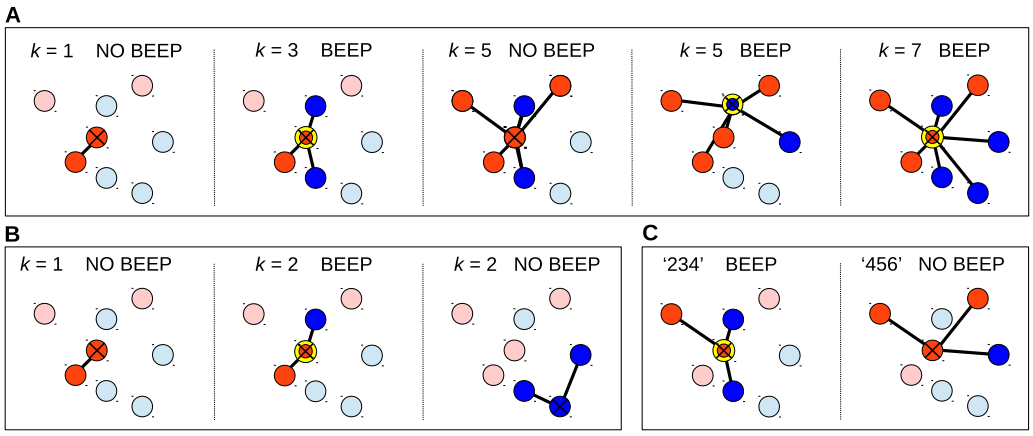

Figure 2.2: Beeping state of a focal individual for the three different criteria used in the segregation experiments. (A): Majoritarian criterion. (B): Exclusive criterion. (C): Shifted environment criterion. Circles with a cross denote focal individuals. Solid lines link the focal individual to neighbors taken into account in each kind of environment (in bright colors). Pale colors denote neighbors not taken into account in the environment of the focal individual. Yellow corona around focal individuals denotes that the tag is beeping. Note that the focal individual is always the same except in the 4^{th} example in (A) and the 2^{nd} example in (B).

example of $k = 5$. The same tag, with the same neighbors, doesn't beep with the exclusive criterion when $k = 1$ (Panel B), but beeps when $k = 2$, while for the shifted environment criterion (Panel C), it beeps in the case '234' (as well as in the case '345' not depicted), but not in the case '456'. Sessions corresponding to different conditions were alternated randomly in order to prevent the participants from detecting any patterns in the sequence of sessions.

Remark: for technical reasons that are still under investigation (probably related to the tags emission frequency which seems to have changed between September 2015 and June 2016), the segregation times were about twice as long in the experiments performed in June 2016 as in the experiments performed in September 2015. The behavior in the random walk phase was of course not affected by this problem because tags and beeps don't play a role in it, so all data from all experiments were used in the construction of the model. In the segregation phase, only the data from September 2015 were used in the graphs related to segregation times. However, the segregation quality (number of clusters formed at the end of the segregation process) was unaffected by this problem, such that all data from both experiments were used in the graphs related to segregation quality. The data used will be explicitely stated in the caption of each figure of the segregation phase.

Instructions given to the participants

In a first series of experiments, called "classic segregation", participants were not explicitly informed that they had to segregate in two clusters. The participants received the following instructions:

1. *Each one of you has been assigned one of two colors, red or blue. None of you knows his own color or the color of others. Your color may change from one session to the next.*
2. *You will first start by walking randomly at a normal pace, while remaining within the limits of the arena.*
3. *After a short transient your tags will be switched on and will start to emit an acoustic signal (a beep) when your environment is not of the same color as you.*

4. The experiment stops when all tags become silent (nobody beeps).

Participants were not allowed to have any oral or facial communication with each other, and were not informed that the two sub-groups were of equal size. Once all the sessions of classic segregation had been carried out, a second series of experiments, called "segregation in 2 clusters", was performed in which the participants were explicitly informed that they had to separate in two clusters. The fourth rule became:

4'. The experiment stops when all tags become silent (nobody beeps), and two clusters are clearly distinguishable.

The "segregation in 2 clusters" experiments were performed after all the "classic segregation" experiments had been carried out because the instruction to "form two clusters" was a key indication that could artificially bias the performance of groups in classic segregation tasks.

2.3 Data collection and preprocessing

Each individual was carrying two tags that sent wireless pulses with a (mean) frequency of 2 Hz, resulting in a stream of location information on the server of about 4 Hz. Wireless signals were scheduled using a time-division scheme in order to avoid collisions, using 134 time slots per second, shared by the 44 tags used in the experiments. Hence, the position of each of the 44 tags was known twice per second at a different instant of time, spanning a time interval of around 1 second. Moreover, due to the time-sharing, tags did not emit signals at exactly 2 Hz, although our system was accurate and the communication flux was highly homogeneous. Fig. 2.3 shows that the disparity within a session is also quite small.

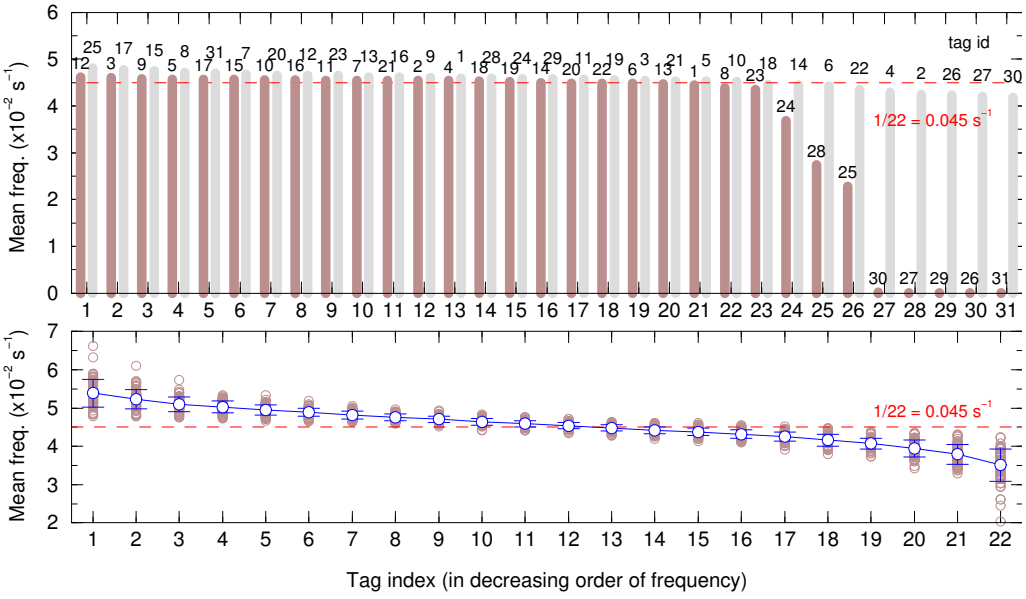

Figure 2.3: Mean frequency of the number of data sent by a tag during a session. Upper panel: mean frequency of each tag, denoted by its identity number. Brown data correspond to September 2015, gray data to June 2016, where more than 30 tags were used during the experiments (due to batteries or individual replacement). Lower panel: mean frequency of the most frequent tag of a session, the second most frequent tag, and so on, until the least frequent tag in the session. The most frequent tag is not always the same from one session to another. The theoretical value is $1/22 = 0.045\ s^{-1}$ *(red dashed lines).*

2.3.1 Structure of data

The data collected by the system for an individual in a given session consists of two series of triplets $\{t_i^{\mathrm{L}}, x_i^{\mathrm{L}}, y_i^{\mathrm{L}}\}_{i=1}^{N_{\mathrm{L}}}$ and $\{t_j^{\mathrm{R}}, x_j^{\mathrm{R}}, y_j^{\mathrm{R}}\}_{j=1}^{N_{\mathrm{R}}}$ corresponding to the left and right tags respectively, where (x_i, y_i) denotes the spatial position of the tag at time t_i. Typically both time series are not of the same length ($N^{\mathrm{L}} \neq N^{\mathrm{R}}$), they are not synchronized ($|t_i^{\mathrm{L}} - t_j^{\mathrm{R}}| > \varepsilon$ for $i = 1, \ldots, N^{\mathrm{L}}$ and $j = 1, \ldots, N^{\mathrm{R}}$, with $\varepsilon \approx 0.2$ s), and although both series are quite regular ($t_{i+1}^{\mathrm{L}} - t_i^{\mathrm{L}} \approx t_{j+1}^{\mathrm{R}} - t_j^{\mathrm{R}} \approx 0.5$ s), there exist some gaps or missing data at different positions in each series. Besides these usual pathologies in data processing, we also detected an artificial temporal non-uniformity in the arrival of data streams, as if data arrived by *waves* (see Fig. 2.4).

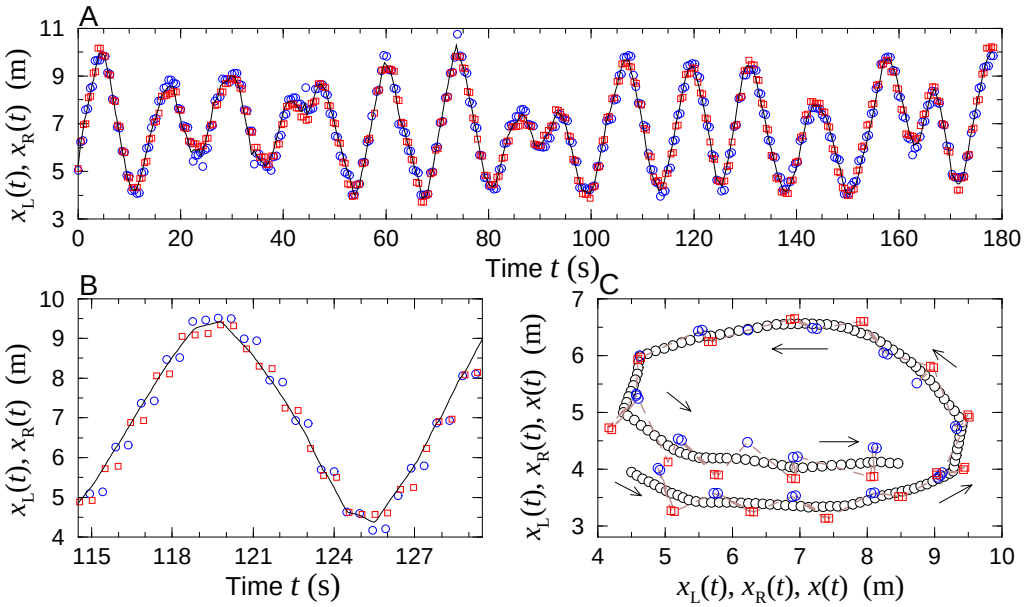

Figure 2.4: Example of a reconstructed trajectory. (A) Time series of the x-coordinates abscissas of the left and right tags of an individual $x_{\mathrm{L}}(t)$ (blue circles) and $x_{\mathrm{R}}(t)$ (red squares) respectively, as extracted from the collected data, together with the equispaced time series (with constant time-step $\Delta t = 0.1$ s) of the reconstructed trajectory's x-coordinate for individual $x_{\mathrm{H}}(s)$. (B) Detail of the time interval $[114.5, 128.5]$ s, showing that tag time series alternate (square–circle), while spatial series of each tag appear by pairs (two squares–two circles), with an alternating spatial gap of size ≈ 0.05 m between spatially consecutive data and of size ≈ 1 m between alternating pairs. (C) Final reconstructed trajectory (black circles) with time step $\Delta s = 0.1$ s corresponding to the time interval shown in (B), and successive position of both tags (joined with a brown dashed line following the time sequence). Arrows denote the anticlockwise direction of motion.

Our guess is that the TCP/IP protocol accumulates the data collected in the router from the different sensors and awaits for data packets to be sufficiently large before transmitting them to the computer. As time is defined by the time of arrival to the database, an artificial delay is generated in the time series, giving rise to an undesirable wave effect in the dynamics of a normal walking pace.

2.3.2 Data preprocessing: synchronization and time rectification

Ideally, the information provided by the two tags of an individual should determine both the position and the orientation of the individual, if it is assumed that the shoulders of the individual are always perpendicular to the direction of motion. This is however not the case: people tend to rotate their body when crossing or approaching each other, or even with no apparent reason. We thus integrated the information of both tags in a single data and consider that individuals

are described by a single point particle located in the geometric center of the segment defined by the two tags. The error in the location of this point is one half of the error in the location of the shoulders. We then considered that the orientation of an individual can be determined by the direction of the velocity vector.

Our analysis required to know the positions of all $N = 22$ individuals simultaneously at each time instant, together with the color of all individuals. Data were synchronized in a common time scale. In order to allow spatial derivation to estimate the velocity and acceleration, we used a small time step of size $\Delta t = 0.1$ s, again by linear interpolation (keeping in mind, during the posterior statistical analysis, that *real* data had a time step of size 0.5 s).

Finally, we corrected the time waves effect by assuming that positions are well calculated and that time steps should simply be redistributed in the time series according to an homogeneous density along the series. The detailed steps of the data reconstruction procedure are as follows:

1. Synchronization of the times series of both tags. Each time series is first completed by linear interpolation in order to have the same instants of time for both tags. Both time series are joined in increasing order in a single time series $\{\hat{t}_i\}_{i=1}^{N_\mathrm{H}}$ of length $N_\mathrm{H} = N_\mathrm{L} + N_\mathrm{R}$. The position of each respective tag at the new instants of time, $\{\hat{x}_i^\mathrm{L}, \hat{y}_i^\mathrm{L}\}_{i=1}^{N_\mathrm{H}}$ and $\{\hat{x}_i^\mathrm{R}, \hat{y}_i^\mathrm{R}\}_{i=1}^{N_\mathrm{H}}$, are calculated by linear interpolation of $\{x_i^\mathrm{L}, y_i^\mathrm{L}\}_{i=1}^{N_\mathrm{L}}$ and $\{x_j^\mathrm{R}, y_j^\mathrm{R}\}_{j=1}^{N_\mathrm{R}}$ respectively, at times $\hat{t}_i$. Then, $\Delta\hat{t}_i = \hat{t}_{i+1} - \hat{t}_i$ is approximately equal to 0.2 and 0.3 s, approximately alternatively.

2. The position of a pedestrian at a given time is then calculated by linear interpolation of the positions of the two tags: $x_i^\mathrm{H} = (\hat{x}_i^\mathrm{L} + \hat{x}_i^\mathrm{R})/2$, $y_i^\mathrm{H} = (\hat{y}_i^\mathrm{L} + \hat{y}_i^\mathrm{R})/2$, for $i = 1, \ldots, N_\mathrm{H}$.

3. Time instants are redistributed according to a locally averaged velocity. The new time series $\{s_i\}_{i=1}^{N_t}$ is built recursively as follows:

$$s_1 = \hat{t}_1, \tag{2.1}$$

$$s_{i+1} = s_i + \lambda \frac{d_i}{\bar{v}_i}, \quad i = 1, \ldots, N_t - 1, \tag{2.2}$$

where $d_i = \|\vec{u}_{i+1} - \vec{u}_i\|$, $\vec{u}_i = (x_i^\mathrm{H}, y_i^\mathrm{H})$, and $\bar{v}_i$ is an averaged velocity,

$$\bar{v}_i = \frac{\displaystyle\sum_{k=-\infty}^{+\infty} \|\vec{u}_{i+1+k} - \vec{u}_{i+k}\|\, e^{-(\hat{t}_{i+k} - \hat{t}_i)^2/t_c^2}}{\displaystyle\sum_{k=-\infty}^{+\infty} (\hat{t}_{i+1+k} - \hat{t}_{i+k})\, e^{-(\hat{t}_{i+k} - \hat{t}_i)^2/t_c^2}}, \tag{2.3}$$

where t_c is the radius of the time interval centred in $\hat{t}_i$ in which the exponential has a significant value and λ is a normalization constant ensuring that the total duration is preserved: $s_N - s_1 = \hat{t}_N - \hat{t}_1$. Then:

$$s_N - s_1 = s_N - s_{N-1} + s_{N-1} - s_{N-2} + s_{N-2} - \cdots - s_2 + s_2 - s_1 \tag{2.4}$$

$$= \lambda \left(\frac{d_{N-1}}{\bar{v}_{N-1}} + \frac{d_{N-2}}{\bar{v}_{N-2}} + \cdots + \frac{d_1}{\bar{v}_1} \right) = \lambda \sum_{i=1}^{N-1} \frac{d_i}{\bar{v}_i} = \hat{t}_N - \hat{t}_1 \tag{2.5}$$

so $\lambda = (\hat{t}_N - \hat{t}_1) \times \left(\displaystyle\sum_{i=1}^{N-1} \frac{d_i}{\bar{v}_i} \right)^{-1}$. Note also that $s_{i+1} = s_1 + \lambda \displaystyle\sum_{j=1}^{i} \frac{d_j}{\bar{v}_j}$, so $s_N = s_1 + \lambda \displaystyle\sum_{i=1}^{N-1} \frac{d_i}{\bar{v}_i}$.

If $t_c = 0$, then only the term with $k = 0$ remains in expression 2.3, and $\bar{v}_i$ becomes

$$\bar{v}_i = \frac{\|\vec{u}_{i+1} - \vec{u}_i\|}{\hat{t}_{i+1} - \hat{t}_i} = \frac{d_i}{\hat{t}_{i+1} - \hat{t}_i}, \tag{2.6}$$

so, for $\lambda = 1$, $s_{i+1} = s_i + \hat{t}_{i+1} - \hat{t}_i$ (*i.e.* $s_i = \hat{t}_i$), so that the time instants are not redistributed.

The critical parameter t_c is the radius of the time window over which the velocity is averaged. The choice of this parameter is based on the compromise of reducing the impact of the bursts on the normal pace while preserving the speed variation proper to normal walk.

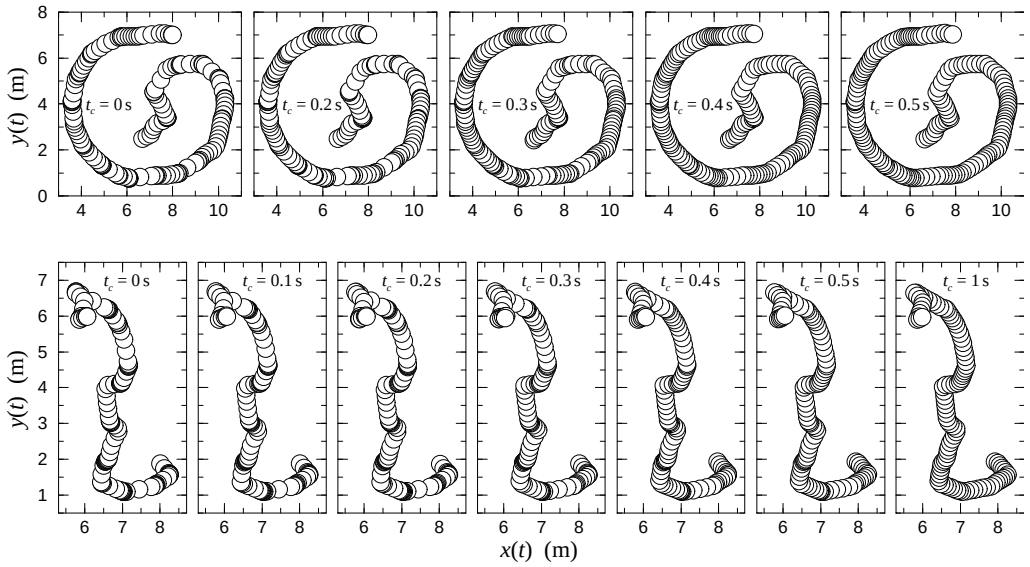

Figure 2.5: Reduction of the burst or wave effect through the redistribution of time instants for increasing values of the averaging parameter t_c. Upper row: Trajectory of a single individual during the random walking phase with the other 21 individuals. Lower row: Same individual, during a segregation task. Circles denote successive positions of the individual at equispaced instants of time ($\Delta t = 0.2$ s). Black rings observed for small values of t_c correspond to the accumulation of circles in short intervals of time, i.e., a burst, or even a stop. The selected value $t_c = 0.4$ s is a compromise between small values for which artificial stops are excessively pronounced and large values for which the excessive smoothing erases the real variations of the speed.

Fig. 2.5 shows the successive positions (circles) of the same individual during a portion of his trajectory when walking randomly (upper panels) and during a segregation task (lower panels). The circles are equispaced in time so that the bursts appear clearly as dark rings, especially for small values of t_c and in the segregation phase, where the speed is in general smaller than in the random walk phase. For the larger values of t_c, the circles are homogeneously redistributed along the trajectory: for $t_c = 0.5$ s in the upper panels and $t_c = 1$ s in the lower ones, the bursts have disappeared, but with the inconvenient loss of diversity in the walking speed, whose variation has been practically removed.

An exhaustive trial and error analysis has shown that the best compromise is reached when $t_c = 0.4$ s. We have also performed an analysis of the spatial error introduced by the redistribution of time instants, by evaluating the spatial shift that individual positions experiment when $t_c > 0$ with respect to their location when $t_c = 0$ s. Note that the position at time t when $t_c = 0$ s is not necessarily the true one, which can only be evaluated by comparing with the videos of the direct experiments, so that this analysis is only a measure of the spatial shift produced by the use of different values of t_c. Figure 2.6 shows that the spatial shift is almost always smaller than 40 cm, which is the typical width of an individual.

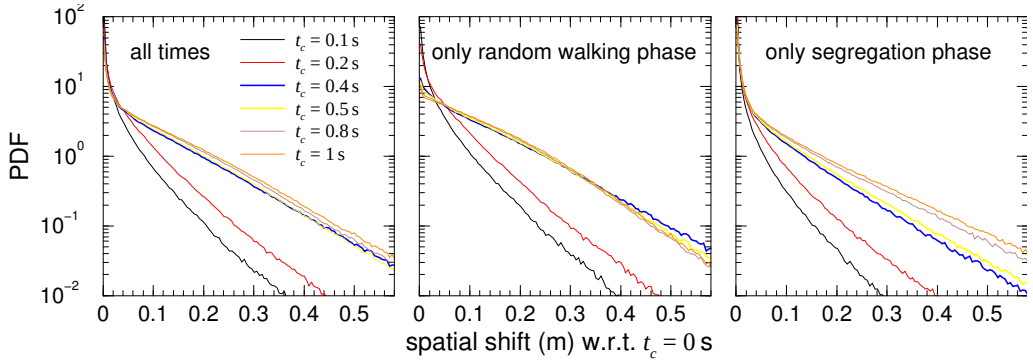

Figure 2.6: Probability density function of the spatial shift of the reconstructed position of an individual at a given time for a value of $t_c > 0$ s, with respect to its position when no reconstruction is carried out ($t_c = 0$ s), for all the individuals of all the sessions with 22 individuals. For small values of t_c the spatial shift is smaller in the segregation phase (around 10 cm smaller) because the speed is generally smaller during this phase. This is not the case for $t_c = 0.8$ and 1 s, where the homogeneisation of speed erases the difference between both phases.

2.4 Dynamics of Pedestrian Motion

2.4.1 Equations of Motion

In continuity of *social force* models, we consider pedestrians as self-driven "particles" subjected to internal "motivations to act" in response to environmental factors, such as the presence of obstacles or other pedestrians [99]. In our case, pedestrians were asked to remain within the limits of a circle drawn on the floor, which constituted a physical border (since it was possible for pedestrians to inadvertently cross it, we will refer to this border as a "soft wall"). They were asked to walk randomly altogether, such that the assumed "forces" at play are a self-propulsion term, a noise term and repulsive interactions between pedestrians and the border as well as between pedestrian themselves. We thus use a Langevin equation of the form:

$$\frac{d\vec{r}_i(t)}{dt} = \vec{v}_i(t), \tag{2.7}$$

$$\frac{d\vec{v}_i(t)}{dt} = -A(\vec{v}_i(t)) + \vec{\eta}_i(t) + \vec{F_{\mathrm{w}_i}}(t) + \sum_{j=1, j\neq i}^{N} \vec{F_{\mathrm{h}_{ij}}}(t), \tag{2.8}$$

where i is the index for individuals (pedestrians), $\vec{r_i}$ and $\vec{v_i}$ are respectively individual i's position (with respect to the center of the arena) and velocity. A is a self-propulsion term, $\vec{\eta}$ a Gaussian white noise, $\vec{F}_{\mathrm{w}}$ and $\vec{F}_{\mathrm{h}}$ respectively the repulsion terms of pedestrian i with the wall and other pedestrians.

To determine the functional form of the different terms, we used an error minimization based procedure (detailed in Appendix 3.3), previously used in [107] to extract the interaction functions of a fish with a wall and between two fish in a circular tank whose dimensions were proportional to the arenas used in the present experiment.

2.4.2 Interaction Forces

More than 250 sessions were performed where individuals were asked to walk randomly at their normal pace. We used these more than 40 hours of data to characterize the motion of pedestrians walking randomly in confined arenas of different sizes. Sessions with a single pedestrian allowed us to measure the self-propulsion force and the interaction with the soft wall, and sessions with two pedestrians the pedestrian-pedestrian interaction force (analysis still in progress). Finally,

sessions with 22 pedestrians were used to compare predictions of the model thus built to collective patterns measured experimentally.

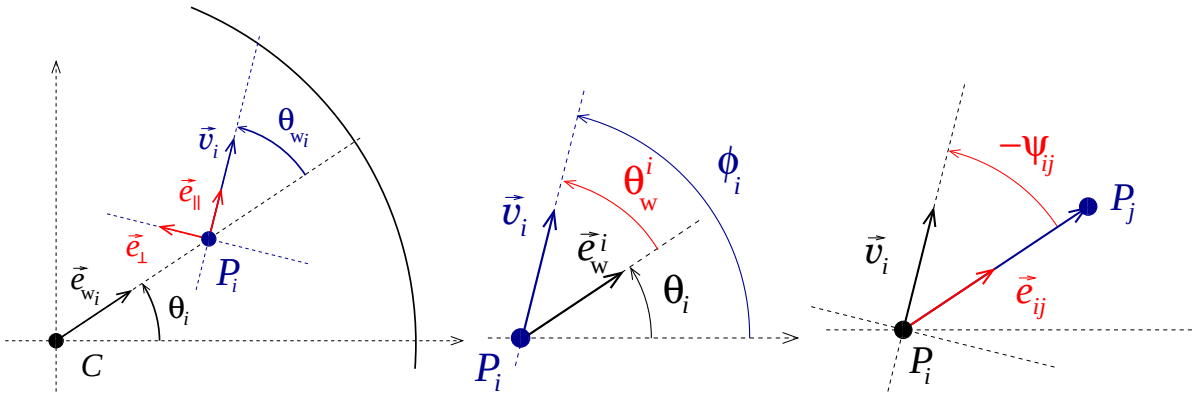

Figure 2.7: Notations used to characterize the position and movement of pedestrians: pedestrian P_i is walking in a circle centered on C, with velocity $\vec{v}_i$ and with an angle θ_{w_i} to the wall. $\vec{e}_{w_i}$ is the unit radial vector, $\vec{e}_{\parallel}$ the unit vector in the direction of motion, $\vec{e}_{\perp}$ the unit vector perpendicular to the direction of motion and $\vec{e}_{ij}$ the unit vector in the direction $\vec{P_iP_j}$ (P_j is another pedestrian). θ_i is the angle between $\vec{e}_{w_i}$ and the horizontal, ϕ_i the angle between the direction of motion and the horizontal and Ψ_{ij} the angle between $\vec{v}_i$ and $\vec{e}_{ij}$.

We now discuss the shape of these interaction functions and the analytical expressions used in the model. Figure 2.7 gives a visual representation of all notations used all along this Chapter and in the Appendix.

Self-Propulsion Term

Figure 2.8 shows that A depends linearly on v:

$$A(\vec{v}) = \frac{v - \bar{v}}{\tau_0}\,\vec{e}_{\parallel} \tag{2.9}$$

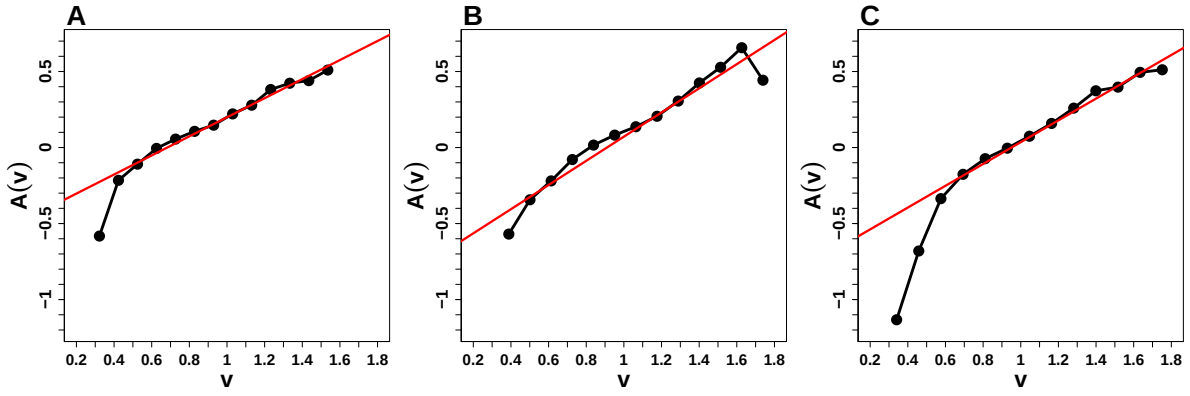

Figure 2.8: In black are the forms of the function $A(v)$ obtained from the data, using the extraction procedure described in Appendix 3.3, for: A. the small arena ($R = 1.78$ m); B. the medium arena ($R = 2.52$ m); C. the large arena ($R = 3.56$ m). In red are linear fits as proposed in the text. All parameter values are given in Table 2.2 at the end of this section.

In the random walk phase, pedestrians must walk continuously. Therefore, we expect them to shortly reach a comfort speed $\bar{v}$, and adjust their speed (norm of the velocity $\vec{v}$) to $\bar{v}$ whenever they are slowed down or sped up (by other pedestrians, the wall, or internal factors). τ_0 is the

typical adjustment time, and $\vec{e_{\parallel}}$ the unit vector in the direction of the velocity (see Figure 2.7). For simplicity, we will assume $\bar{v}$ to be the same for all pedestrians, as further analysis showed that choosing $\bar{v}$ for each individual according to the actual distribution of "intrinsic" speeds (defined as the total distance spanned divided by the corresponding time spent walking) doesn't change the results significantly.

Noise Term

The white Gaussian noise $\vec{\eta}$ can be understood as a pedestrian's "free will", in the sense of the resultant of all internal factors that motivate her to move in a certain direction with a certain speed. By analogy with the Brownian motion of actual particles in a fluid, we define:

$$\vec{\eta}(t) = v_0 \sqrt{\frac{2}{\tau_0}} \begin{cases} g_x(t) \\ g_y(t) \end{cases} \tag{2.10}$$

with $\langle g_\alpha(t) \rangle = 0, \langle g_\alpha(t) \cdot g_\alpha(t') \rangle = \delta(t - t')$, where $\alpha = x$ or y, t and t' are two different times in the process, v_0 is the typical speed of pedestrians and τ_0 is the adjustment time above defined.

Note that the reconstruction procedure also gave us the actual functional forms of the parallel and perpendicular (to the direction of motion) components of the noise, which correspond to more complicated analytical expressions. Yet, for isotropy reasons, the simpler expressions proposed in equation 2.10 yield the same collective outcomes. For the sake of simplicity, we thus chose to keep them.

Interaction with the Soft Wall

Circular arenas can be considered an isotropic medium (no privileged direction or position), so that $\vec{F}_{\mathrm{w}}(t)$ can be taken as directed towards the center of the arena (centripetal).

Let $\theta_{\mathrm{w}}(t)$ the angle of incidence of the velocity vector with respect to the border of the arena (see Figure 2.7): $\theta_{\mathrm{w}}(t) = 0$ means that the pedestrian is walking towards the wall, $\theta_{\mathrm{w}}(t) = \pm\pi$ means that the pedestrian is walking towards the center of the arena. The angle $\theta_{\mathrm{w}}(t)$ is defined positive (resp. negative) when the pedestrian has the wall at her right (resp. left).

We expect the magnitude of the force $\vec{F}_{\mathrm{w}}$ to be determined by the distance $r_{\mathrm{w}}(t) = R - r(t)$ (where $r = |\vec{r}|$) between a pedestrian and the wall, as well as by his angle $\theta_{\mathrm{w}}(t)$ and speed $v(t)$ of incidence to the wall. Assuming that the aforementioned contributions are decoupled from one another, the force can be written:

$$\vec{F}_{\mathrm{w}}(v, r_{\mathrm{w}}, \theta_{\mathrm{w}}) = -B_{\mathrm{w}}(v) f_{\mathrm{w}}(r_{\mathrm{w}}) g_{\mathrm{w}}(\theta_{\mathrm{w}}) \vec{e}_{\mathrm{w}}, \tag{2.11}$$

where $\vec{e}_{\mathrm{w}}$ is the unit radial vector (see Figure 2.7). Further analysis showed that $B_{\mathrm{w}}(v)$ can be considered equal to one (no dependence on speed), and from Figures 2.9 and 2.10 we deduce:

$$g_{\mathrm{w}}(\theta_{\mathrm{w}}) = a_{\mathrm{w}_0} + a_{\mathrm{w}_1} \cos(\theta_{\mathrm{w}}) + a_{\mathrm{w}_2} \cos(2\theta_{\mathrm{w}}) + a_{\mathrm{w}_3} \cos(3\theta_{\mathrm{w}}) + a_{\mathrm{w}_4} \cos(4\theta_{\mathrm{w}}) \tag{2.12}$$

$$f_{\mathrm{w}}(r_{\mathrm{w}}) = \begin{cases} a_{\mathrm{w}} \left(e^{\frac{-r_{\mathrm{w}}}{l_{\mathrm{w}}}} - e^{\frac{-r_{\mathrm{w_c}}}{l_{\mathrm{w}}}} \right) & \text{if } r_{\mathrm{w}} < r_{\mathrm{w_c}} \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

where $r_{\mathrm{w_c}}$ is the critical distance to the wall beyond which pedestrians don't "feel" the wall anymore ($f_{\mathrm{w}}(r_{\mathrm{w}}) = 0$), and l_{w} is the typical interaction range. The a_{w}s are coefficients which values are given in Table 2.2.

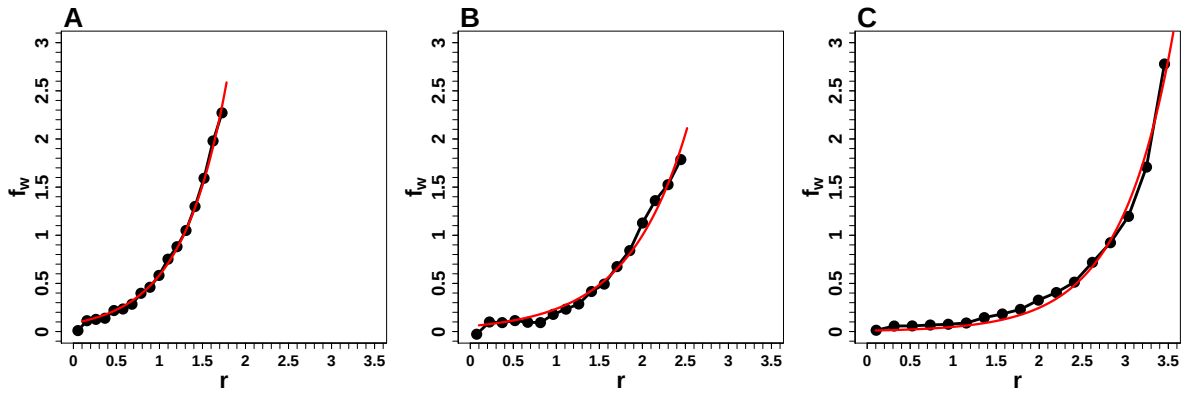

Figure 2.9: In black are the forms of the radial component $f_w(r)$ obtained from the data, using the extraction procedure described in Appendix 3.3, for: A. the small arena ($R = 1.78$ m); B. the medium arena ($R = 2.52$ m); C. the large arena ($R = 3.56$ m). In red are exponential fits as proposed in the text. All parameter values are given in Table 2.2 at the end of this section.

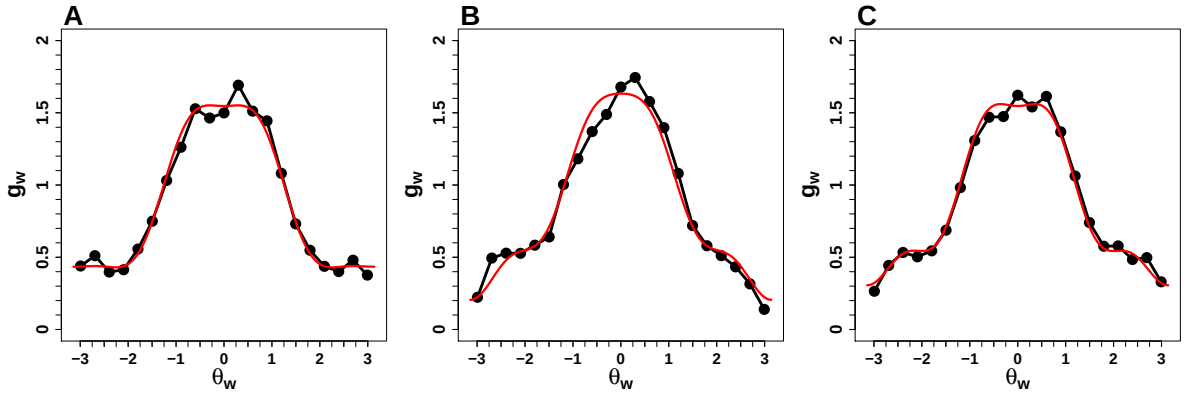

Figure 2.10: In black are the forms of the angular component $g_w(\theta_w)$ of the interaction with the wall, as functions of the angle θ_w to the wall, obtained from the data, using the extraction procedure described in Appendix 3.3, for: A. the small arena ($R = 1.78$ m); B. the medium arena ($R = 2.52$ m); C. the large arena ($R = 3.56$ m). In red are cosine fits as proposed in the text. All parameter values are given in Table 2.2 at the end of this section.

Interaction Between Pedestrians

Similarly, we assume (the extraction of the actual form of the interaction being still in progress) that the force exerted by pedestrian P_j on pedestrian P_i ($j \neq i$) depends on the distance $r_{ij} = |\vec{r}_i - \vec{r}_j|$ between them, on their relative speed $v_{ij} = |\vec{v}_i - \vec{v}_j|$ and on the angle $\psi_{ij}(t)$ (in general, $\psi_{ij}(t) \neq \psi_{ji}(t)$) with which pedestrian P_i perceives pedestrian P_j (see Figure 2.7):

$$\vec{F}_{\mathrm{h}_{ij}}(v_{ij}, r_{ij}, \psi_{ij}) = -B_{\mathrm{h}}(v_{ij}) f_{\mathrm{h}}(r_{ij}) g_{\mathrm{h}}(\psi_{ij}) \vec{e}_{ij}, \tag{2.14}$$

with $\vec{e}_{ij}$ a unit vector in the direction P_iP_j (see Figure 2.7). We use the same analytical expression for g_{h} as for g_{w}: $g_{\mathrm{h}}(\psi_{ij}) = g_{\mathrm{w}}(\psi_{ij})$, neglect the effect of speed as previously ($B_{\mathrm{h}}(v_{ij}) = 1$), and assume a similar form for $f_{\mathrm{h}}(r_{ij})$, but with a stronger repulsion at short distance (see Figure 2.11):

$$f_{\mathrm{h}}(r_{ij}) = \begin{cases} a_{\mathrm{h}} \left(e^{-\left(\frac{r_{ij}}{l_{\mathrm{h}}}\right)^2} - e^{-\left(\frac{r_{\mathrm{h_c}}}{l_{\mathrm{h}}}\right)^2} \right) & \text{if } r_{ij} < r_{\mathrm{h_c}} \\ 0 & \text{otherwise} \end{cases} \tag{2.15}$$

where $r_{\mathrm{h_c}}$ is the critical distance between individuals, beyond which the interaction recedes ($f_{\mathrm{h}}(r_{ij}) = 0$), and l_{h} is the typical interaction range.

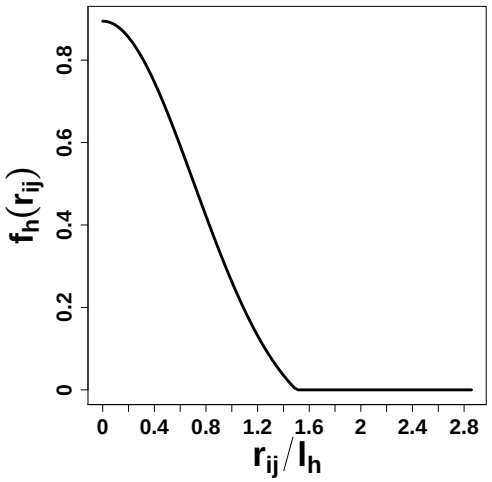

Figure 2.11: Radial dependence f_h of the repulsion force between pedestrians. All parameter values are given in Table 2.2 at the end of this section.

Numerical Approximation

Let $h = 10^{-2}$ s be the integration step[3]. We know that in the Ornstein-Uhlenbeck process, the noise term is proportional to $\sqrt{h}$. We assume h small enough for the forces to be constant in such a time interval, and we compute, at each time step t, the forces at $t + \frac{h}{2}$. Therefore, equation 2.8 numerically reads, for each component α ($\alpha = x$, y):

$$v_\alpha(t+h) = v_\alpha(t) - h\,A(v_\alpha(t)) + \sqrt{h}\,v_0\sqrt{\frac{2}{\tau_0}}\,g_\alpha(t) + h\,F_\alpha(t+\frac{h}{2}) \tag{2.16}$$

where $F = F_\mathrm{w} + F_\mathrm{h}$. We approximate r_α at order 2:

$$r_\alpha(t+h) = r_\alpha(t) + h\,v_\alpha(t) + \frac{h^2}{2}\frac{v_\alpha(t+h) - v_\alpha(t)}{h} \tag{2.17}$$

$$= r_\alpha(t) + \frac{h}{2}(v_\alpha(t+h) + v_\alpha(t)) \tag{2.18}$$

We now apply our model of pedestrian motion on groups of 22 pedestrians, and compare correlation functions and various distributions of collective motion to experimental measures.

2.4.3 Correlation Functions

We define the speed autocorrelation and average squared distance as:

$$C(t) = \langle \vec{v}_{e,i}(t'+t) \cdot \vec{v}_{e,i}(t') \rangle_{e,i,t'} \tag{2.19}$$

$$\chi(t) = \langle (\vec{r}_{e,i}(t'+t) - \vec{r}_{e,i}(t'))^2 \rangle_{e,i,t'} \quad , \tag{2.20}$$

where e, i and t' are respectively the indexes running on N_e experiments, N_i individuals and $N_{t'}$ discrete time steps. Note that $N_{t'} = N_{t_e} - t$, where N_{t_e} is the total number of time steps in experiment e. In other words, for $t = 0$, t' runs through all time steps, while for $t = N_{t_e}$, t' can only take the value 0.

[3]We checked that our simulations are stable for time steps smaller or equal to 10^{-2} s.

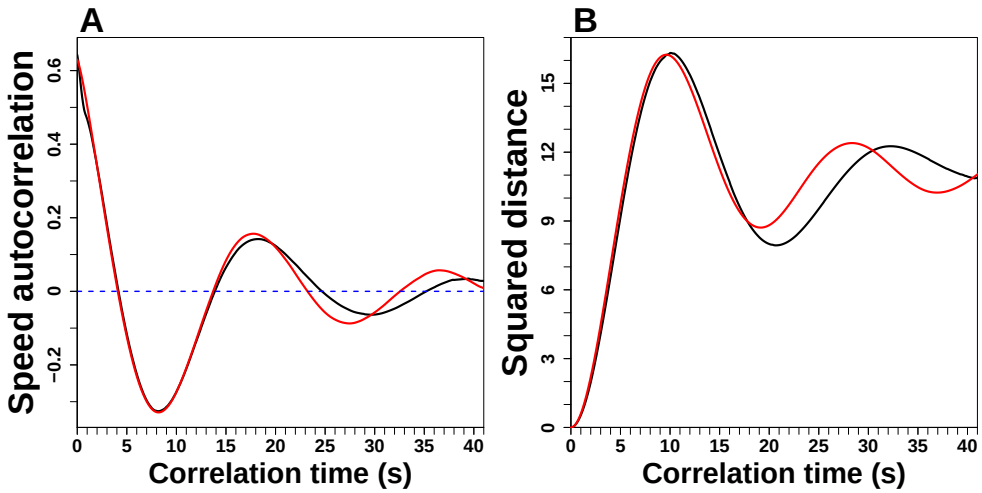

Figure 2.12: A. Speed autocorrelation function $C(t) = \langle \vec{v}_{e,i}(t' + t) \cdot \vec{v}_{e,i}(t') \rangle_{e,i,t'}$*; B. Average squared distance* $\chi(t) = \langle (\vec{r}_{e,i}(t' + t) - \vec{r}_{e,i}(t'))^2 \rangle_{e,i,t'}$*, against the correlation time. The data are in black and the model simulations are in red.*

Speed Autocorrelation Function

As indicated by its name, C is a measure of the extent to which speeds are correlated in time (see Figure 2.12A). One naturally expects that C should decrease as t' increases (*i.e.* as the distance in time between two measures of speed increases), and progressively tend to 0.

Interestingly, we observe an oscillating pattern, with negative values of C suggesting that speeds anti-correlate, and then correlate again, with a periodicity of about 8 to 9 s: because pedestrians are confined in a circular (and thus bounded) arena, their regular encounter with the wall forces them to turn around, thus creating anti-correlations (correlations with opposite speed directions).

Diffusion Process

The average squared distance χ is a usual representation of diffusion processes such as Brownian motion and Ornstein-Uhlenbeck processes (see Figure 2.12B). It measures the extent to which "particles" (pedestrians in the present case) diffuse in space (the distance between two points in time increases on average).

Similarly to the pattern observed in the speed autocorrelation function, the movement of pedestrians being restricted by the circular arena, the diffusion tends to a finite value with oscillatory patterns, also due to encounters with the wall.

Adjustment of Parameters

The noise parameter v_0 and the speed parameter $\bar{v}$ strongly affect the magnitude of oscillations. Indeed, the effect of a stronger noise is to further break the patterns and regularities, hence weakening the correlations and making them vanish faster.

Similarly, with lower comfort speed $\bar{v}$, pedestrians would take more time to reach the wall, leaving them also more time to interact with other pedestrians, which would in turn have a similar effect to more noise.

τ_0 has a weaker but non-negligible impact: higher values of τ_0 increase the pseudo-period and decrease the amplitude of oscillations.

2.4.4 Distributions

Figure 2.13 shows the distributions of speeds, positions, distances to the closest neighbor and angles to the wall. These measures allowed us to finely tune the parameter values.

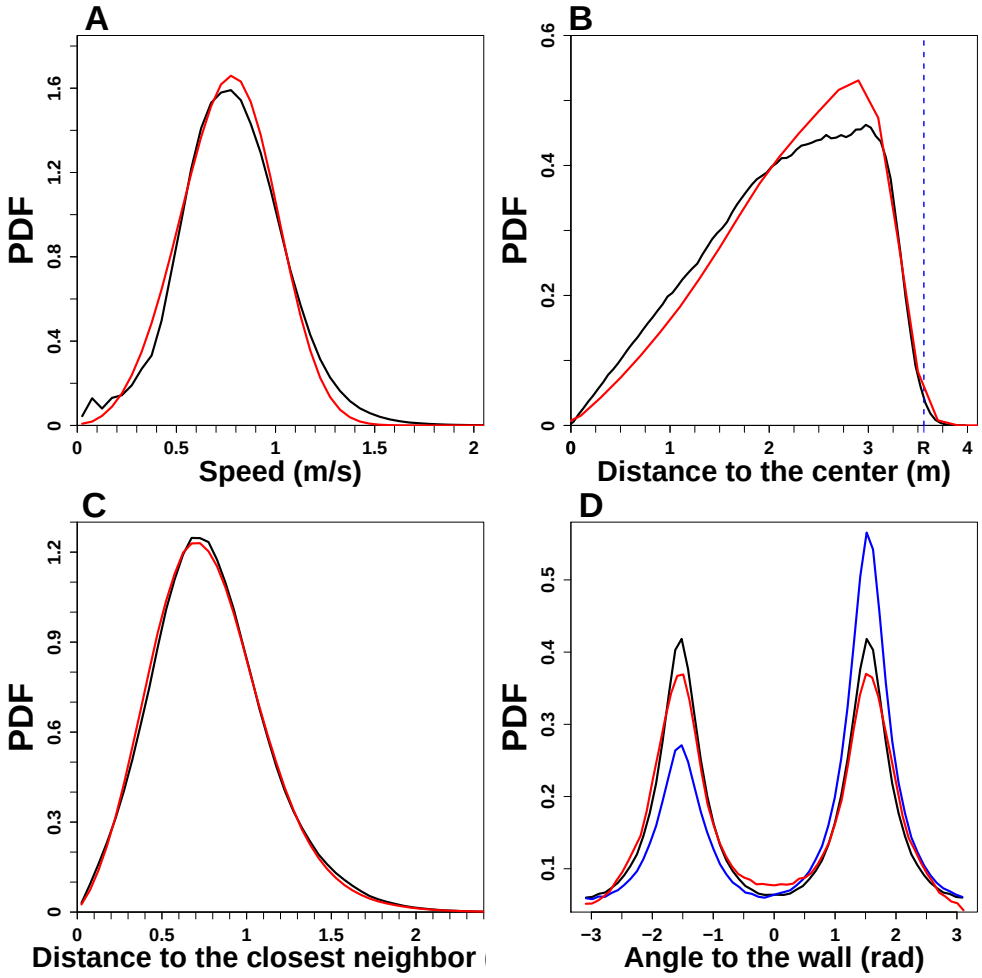

Figure 2.13: A. Distribution of speeds; B. Distribution of positions (distances to the center); C. Distribution of distances to the closest neighbor; The data are in black and the model simulation is in red. D. Distribution of angles to the wall. The original data are in blue, and are not symmetric: pedestrians prefer keeping the wall to their right. In black are the symmetrized data (for each angle we added the opposite angle), and in red is the simulation from the model.

Distribution of speeds

Figure 2.13A shows that speeds are close to Gaussian distributed, and centered at $\sim$0.7 m/s. $\bar{v}$ controls the position of the distribution center, while v_0 influences its width (the higher the noise, the wider the distribution). Higher values of τ_0 also widen the distribution (pedestrians take more time to adjust their speed to $\bar{v}$, and thus the variability in v is higher), and at the same time reduce the average speed (the center shifts to the left).

Distribution of positions

Figure 2.13B shows the distribution of positions of pedestrians. At constant density we would expect the proportion of individuals at a certain distance r from the center to grow linearly with r, which is what we observe for $r < 2$ m.

Above 2 m though, we see the effect of the wall repulsion ($l_{\mathrm{w}} \approx 1$ m): pedestrians try to avoid it, so their proportion decreases rapidly. Note that the proportion of individuals at distance $r \geq R$ from the center is not null, which can be accounted for by the precisions of the measures ($\lesssim 30$ cm). We can't exclude the possibility that individuals sometimes slightly crossed the border (hence the notion of "soft wall").

The interaction magnitude a_{w} and critical interaction distance $R_{\mathrm{w_c}}$ control the shape of the distribution: the higher a_{w}, the stronger the wall repulsion, and the more the distribution is shifted to the left (far from the wall); the higher $R_{\mathrm{w_c}}$, the farther away from the wall would individuals "feel" the wall, and hence the more left-shifted the distribution.

Distribution of distances to the closest neighbor

The distribution of distances to the closest neighbor (Figure 2.13C) is slightly right-skewed and peaked around 0.6 m. The interpretation of parameters l_{h}, a_{h} and $R_{\mathrm{h_c}}$ is similar to their counterparts in the interaction with the wall: l_{h} is the typical distance at which pedestrians start to avoid each other; $R_{\mathrm{h_c}}$ is the distance above which the interaction vanishes, and a_{h} is the intensity of the repulsion between individuals. l_{h} sets the center of the distribution, while increasing values of a_{h} and $R_{\mathrm{h_c}}$ "push" the distribution rightwards, as the repulsion between individuals increases.

Distribution of angles to the wall

The two peaks in Figure 2.13D at $\pm\frac{\pi}{2}$ suggest that pedestrians have a high tendency to walk parallel to the wall, namely to turn around the arena. The function $g_{\mathrm{w}}(\theta_{\mathrm{w}})$ is essential to reproduce this feature, and a similar dependence on θ_{w} in the interaction with a circular border was observed in fish [88, 107]. The left-right asymmetry (blue line), showing that (French) pedestrians prefer to keep the wall on their right, is reminiscent of cultural biases in the choices of the side avoidance of obstacles [103].

To avoid taking into account the left/right asymmetry in our model (which brings nothing to the issues tackled in this chapter), we symmetrized the data (black and red lines) by considering for each angle the opposite angle. It amounts to consider that the real and reversed trajectories are equivalent (like looking at the motion from below or above). To conclude this section, Table 2.2 gives all the parameter values for the random walk phase.

Parameter	Value	Parameter	value
τ_0	0.9	$R_{\mathrm{w_c}}$	$2.5\,l_{\mathrm{w}}$
v_0	0.16	$R_{\mathrm{h_c}}$	$1.6\,l_{\mathrm{h}}$
$\bar{v}$	1.025	$a_{\mathrm{w_0}}$	0.86
a_{w}	1.23	$a_{\mathrm{w_1}}$	0.69
a_{h}	0.9	$a_{\mathrm{w_2}}$	0.0025
l_{w}	1.4	$a_{\mathrm{w_3}}$	-0.018
l_{h}	1.0	$a_{\mathrm{w_4}}$	-0.14

Table 2.2: Parameter values for the random walk phase.

2.5 Phase Separation under Controlled Filtered Information

The second and main objective of this research was to look at how one could influence collective pedestrian behavior using an artificial sensory device able to reduce complex environmental information (*i.e.* positions and colors of some neighbors, themselves selected according to non-trivial rules) into a bit of information easily processable by human beings. In particular, we were interested in the impact of the amount of information processed by the sensory device on phase separation time and quality, defined as the number of distinct clusters at final time.

2.5.1 Model Extension

To describe and analyse the segregation phase, we used the same model of pedestrian motion built and calibrated for the random walk phase, and added a component describing the effects of the device, namely a parameter $\varepsilon_i(t)$ which equals 1 when individual i is beeping at time t, and 0 otherwise. The determination of beeps can obey any kind of rule. When an individual beeps, his motion just follows equation 2.8, and when he stops beeping, we replace the noise and self-propulsion terms by a *friction term* representing his "motivation" to stop. The equation of motion thus reads:

$$\frac{d\vec{v}_i(t)}{dt} = (1-\varepsilon_i(t))\left(-\frac{\vec{v}_i(t)}{\tau}\right) + \varepsilon_i(t)\,(-A(\vec{v}_i(t)) + \vec{\eta}_i(t)) + \vec{F_{\mathrm{w}_i}}(t) + \sum_{j=1,\,j\neq i}^{N} \vec{F_{\mathrm{h}_{ij}}}(t) \qquad (2.21)$$

New parameter values

The dynamics changes in the segregation phase. In particular, the average speed decreases because pedestrians are focusing on finding environments where they stop beeping. They also "accept" to be much closer to each other (l_{h} decreases) once they find members of the same group. The typical stopping time τ is quite large (about 3 s), because the Ubisense tags emitted about one beep per second, such that pedestrians needed at least 2 s to make sure the beep had stopped. Also, we found that pedestrians often form clusters close to the border, such that the range of interaction with the wall decreases (l_{w} and $R_{\mathrm{w_c}}$ decrease). Table 2.3 gives the values of the parameters in the segregation phase.

Parameter	Value	Parameter	value
τ	3	τ_0	5
$\bar{v}$	0.4	v_0	0.18
l_{w}	0.5	l_{h}	0.2
$R_{\mathrm{w_c}}$	l_{w}	$R_{\mathrm{h_c}}$	l_{h}
a_{w}	2.8	a_{h}	1.5

Table 2.3: Parameter values fitted for the segregation phase. The coefficients a_{w_i} in the function g_{w} keep the same values as in the random walk phase.

2.5.2 Time Evolution of the Fraction of Beeps and Number of Clusters

Figures 2.14 and 2.15 show the time evolution of the fraction of subjects beeping and the average number of clusters, respectively. $k = 1$ is the noisiest case: every time pedestrians of different colors come close, they start beeping. This instability makes the segregation time a bit longer in our model than in the data (Figures 2.14A and 2.15A), suggesting that in reality pedestrians adapt their strategy: it is likely that when a pedestrian starts beeping after having been at rest for a certain amount of time (because she is in a "good" environment, *i.e.* with peers of the same color), she sort of guesses that the newcomer (pedestrian of the other color making her

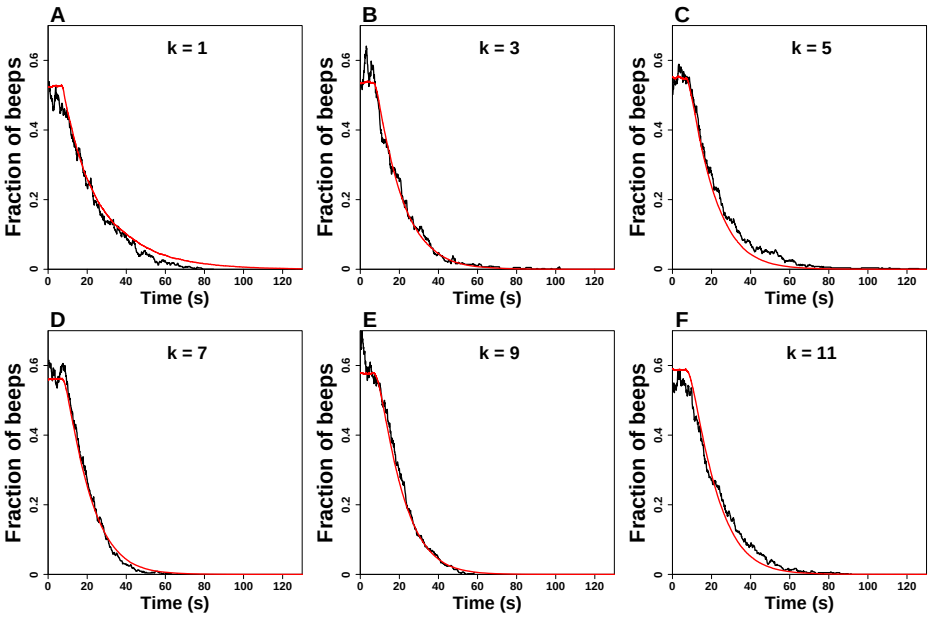

Figure 2.14: Time evolution of the fraction of subjects beeping, for each value of k. Only the data from September are considered here, which includes values of k up to 11. The data are in black, while the simulations from the model are in red.

beep) is "wrong" rather than her, such that she doesn't move. We checked that this mechanism can actually explain the shorter times observed in the data for $k = 1$. However, this effect is small enough to be considered negligible.

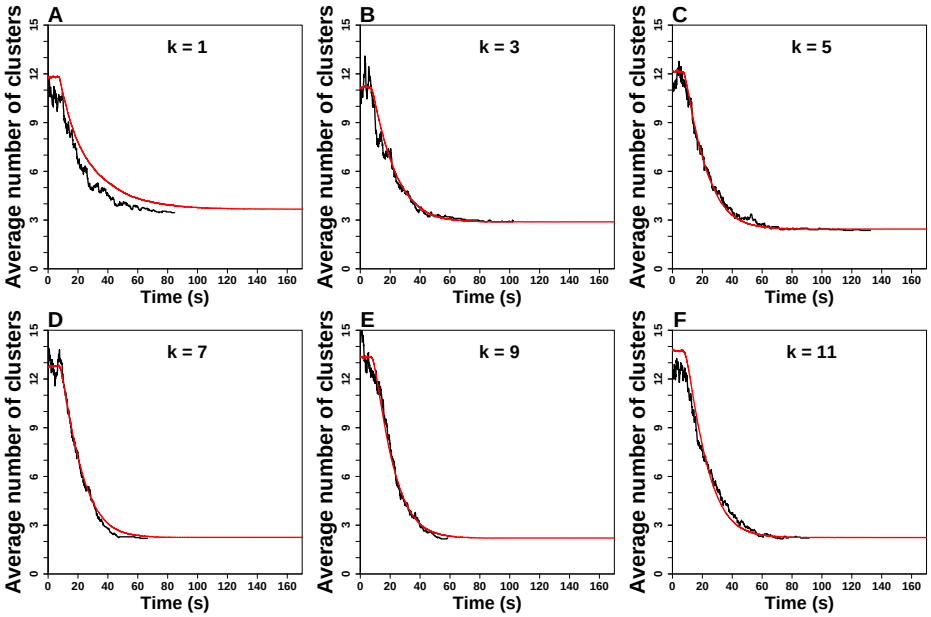

Figure 2.15: Time evolution of the average number of clusters, for each value of k. Only the data from September are considered here, which includes values of k up to 11. The data are in black, while the simulations from the model are in red.

Remark that the decrease doesn't start immediately, suggesting an adaptation time to the transition between random walk and segregation phases. To account for it in the model, we let agents continue as in the random walk phase ($\varepsilon_i = 1 \, \forall i$) for 7 s after the segregation phase begins.

These figures give a general idea of the segregation dynamics. Both segregation time and average number of clusters at final time seem to depend on k. In the next section, we define quantifiers of segregation time and analyse how they are impacted by k.

2.5.3 Impact of Information Quantity on Phase Separation Time

The segregation time can be defined in several manners. We propose three natural measures:

- $\langle t_{\mathrm{b}_{i,e}} \rangle_{i,e}$, where $t_{\mathrm{b}_{i,e}}$ is the total time individual i in experiment e has spent beeping,
- $\langle t_{\mathrm{f}_{i,e}} \rangle_{i,e}$, where $t_{\mathrm{f}_{i,e}}$ is the last time individual i in experiment e has beeped ($t_{\mathrm{f}_{i,e}} \geq t_{\mathrm{b}_{i,e}}$),
- $\langle t_{\mathrm{end}_e} \rangle_e$, where $t_{\mathrm{end}_e} = \mathrm{Max}_i(t_{\mathrm{f}_{i,e}})$ is the last beeping time in experiment e.

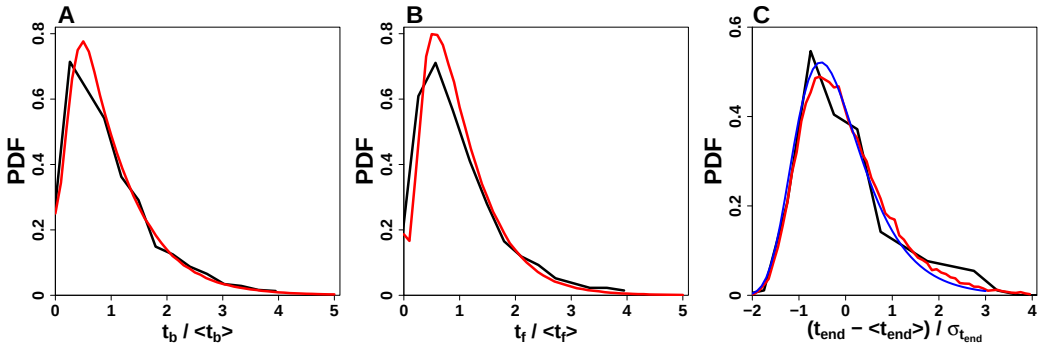

Figure 2.16: Distribution of the normalized segregation time defined as A. t_{b}: the total time an individual spends beeping; B. t_{f}: the last time an individual has beeped; C. t_{end}: the maximum value of t_{f} for an experiment. The data are in black while the simulations from the model are in red. The normalization allowed to combine all data for both experiments in September 2015 and June 2016. In C. remark that t_{end} is very close to a Gumbel distribution (in blue). All values of k are combined.

Figure 2.16 shows the distributions of these quantities, normalized so as to be able to combine values for all k. Remark that the normalized t_{end} follow a Gumbel distribution, which is the distribution of maximum events ($t_{\mathrm{end}} = \mathrm{Max}_i(t_{\mathrm{f}_i})$) of a number of samples, for various distributions.

Figure 2.17 shows the dependence on k of the segregation time for the three selected measures. The data show an optimum value of k at $k = 3$ for $\langle t_b \rangle$ (Figure 2.17A), and around $k = 7 \sim 9$ for $\langle t_f \rangle$ (Figure 2.17B) and $\langle t_{\mathrm{end}} \rangle$ (Figure 2.17C). Our model reproduces well the effect of k on $\langle t_b \rangle$ and $\langle t_f \rangle$ but predicts a weaker impact of k on $\langle t_{\mathrm{end}} \rangle$, especially for $k = 7 \sim 9$.

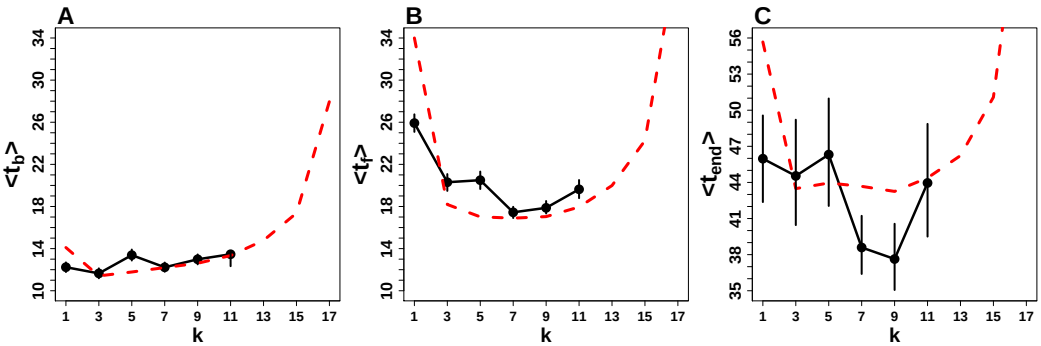

Figure 2.17: Segregation time defined as A. t_{b}: the total time an individual spends beeping; B. t_{f}: the last time an individual has beeped; C. t_{end}: the maximum value of t_{f} for an experiment, against the information parameter k. The data are in black while the simulations from the model are in red. Only the data from the experiment performed in September 2015 were used.

Our model predicts that the segregation time increases drastically for $k > 11$, namely when k is higher than half the group size ($N = 22$). In particular, it would be infinite for $k = 21$, as in such a case the majority of a pedestrian's neighbors would necessarily be of the other color, and thus all pedestrians would ceaselessly beep. Model predictions for different group sizes will be given in section 2.5.5.

2.5.4 Impact of Information Quantity on Phase Separation Quality

Figure 2.18 shows the distribution of cluster sizes at the end of the segregation process. As k increases, the proportion of clusters of size 11 increases up to $k = 9$, and then stabilizes. Our definition of clusters is based on local 3-clusters: a given pedestrian is "linked" to those of his 3 closest neighbors that are of the same color as him. 3-clusters yielded a better agreement than k-clusters (same definition but with k neighbors, k being the information parameter) with the actual group sizes observed (visually) at the end of each experiment, and also have the advantage to be the same basis for all values of k (all experiments). We checked that using k-clusters instead doesn't change any of the conclusions drawn in this section.

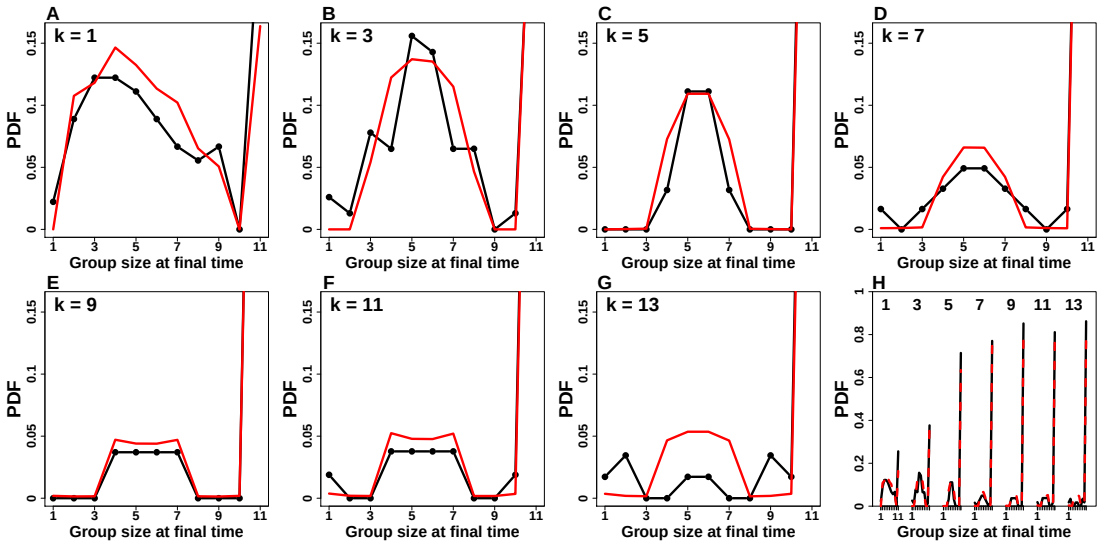

Figure 2.18: Distribution of cluster sizes for $k = 1$ to 13 (A to G) at final time. Since the proportion of clusters of size 11 is much higher than the other cluster sizes, we show it separately in H., with all distributions next to one another. The data are in black and the model simulations are in red. Data from September 2015 and June 2016 were used.

Similarly to what we did in the last section, we define several measures of segregation quality: the average cluster size, the average number of clusters, 22 divided by the average number of clusters (similar but not identical to the average cluster size), the fraction of clusters of size 11, the fraction of "perfect" segregations and the fraction of experiments with at least one cluster of size 11 at final time. By analogy with phase separation, we qualify a segregation process as "better" than another if less (and larger sized) clusters are formed at the end of the process. Following this definition, a segregation process is "perfect" when only two clusters (of size 11) remain at final time (the two "phases" are fully separated).

Figure 2.19 shows the dependence on k of all the above defined segregation quality criteria: segregation quality improves up to $k = 9$, and then saturates. Given that $k = 9$ also optimized segregation time, but since this value is very close to half the group size (11), we need to investigate how this value would change (or not) with different group sizes.

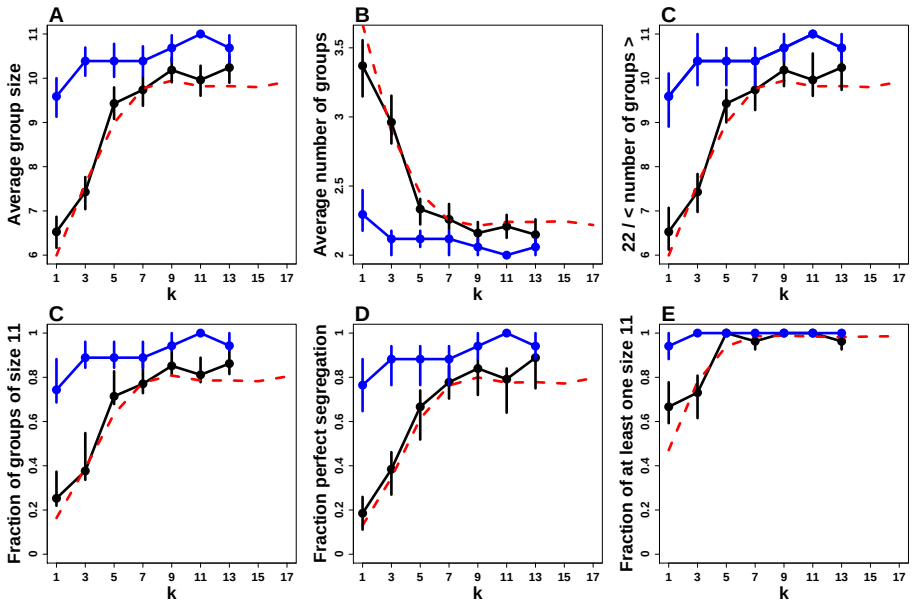

Figure 2.19: A. Average cluster size; B. Average number of clusters; C. 22 / Average number of clusters; D. Fraction of clusters of size 11; E. Fraction of experiments with perfect segregation (2 clusters of size 11); F. Fraction of experiments with at least one cluster of size 11 at final time, against the information parameter k. The data are in black and the model simulations are in red. In blue are the experimental data for the condition when subjects were asked to segregate in two clusters (2-Clusters condition). Data from September 2015 and June 2016 were used.

Remark: notice that when subjects were explicitly asked to form 2 clusters at the end of the process, the segregation quality improved dramatically, especially for low values of k, but yet not to 100%.

2.5.5 Model Predictions for Different Group Sizes

We used the model to analyse the impact of k on segregation time (Figure 2.20) and quality (Figure 2.21) for groups of $N = 10$, 22, 30 and 42 pedestrians, at constant density (the radius of the arena was changed accordingly). The numbers were chosen such that half the group size be an odd number, as all values of k (which were so to have unambiguous definitions of majority). That is why we chose 42 instead of 40, and 22 instead of 20 (same results as previously; we show them again for comparison).

Interestingly, the optimum value of the segregation time (see Figure 2.20) is independent of group size: $k = 3$ for $\langle t_b \rangle$ and $\langle t_{\text{end}} \rangle$ (except $N = 22$ where there is a "flat" minimum from $k = 3$ to $k = 9$), and at $k = 5 \sim 7$ for t_f. Overall, except for the case $k = 1$ which is too noisy, small values of k have to be preferred to minimize the segregation time.

As for segregation quality (see Figure 2.21), although the general patterns are the same for all group sizes, one observes noticeable differences:

- For $N = 10$, pedestrians segregate almost always perfectly (two clusters of 5 individuals) for $k = 1$, and always perfectly for $k \geq 3$.
- For $N \geq 22$, the plateau (optimum segregation quality) is reached for increasingly smaller values of k as compared to half the group size: $k = 7 \sim 9$ for $\frac{N}{2} = 11$, $k = 9 \sim 11$ for $\frac{N}{2} = 15$ and $k \sim 13$ for $\frac{N}{2} = 21$.

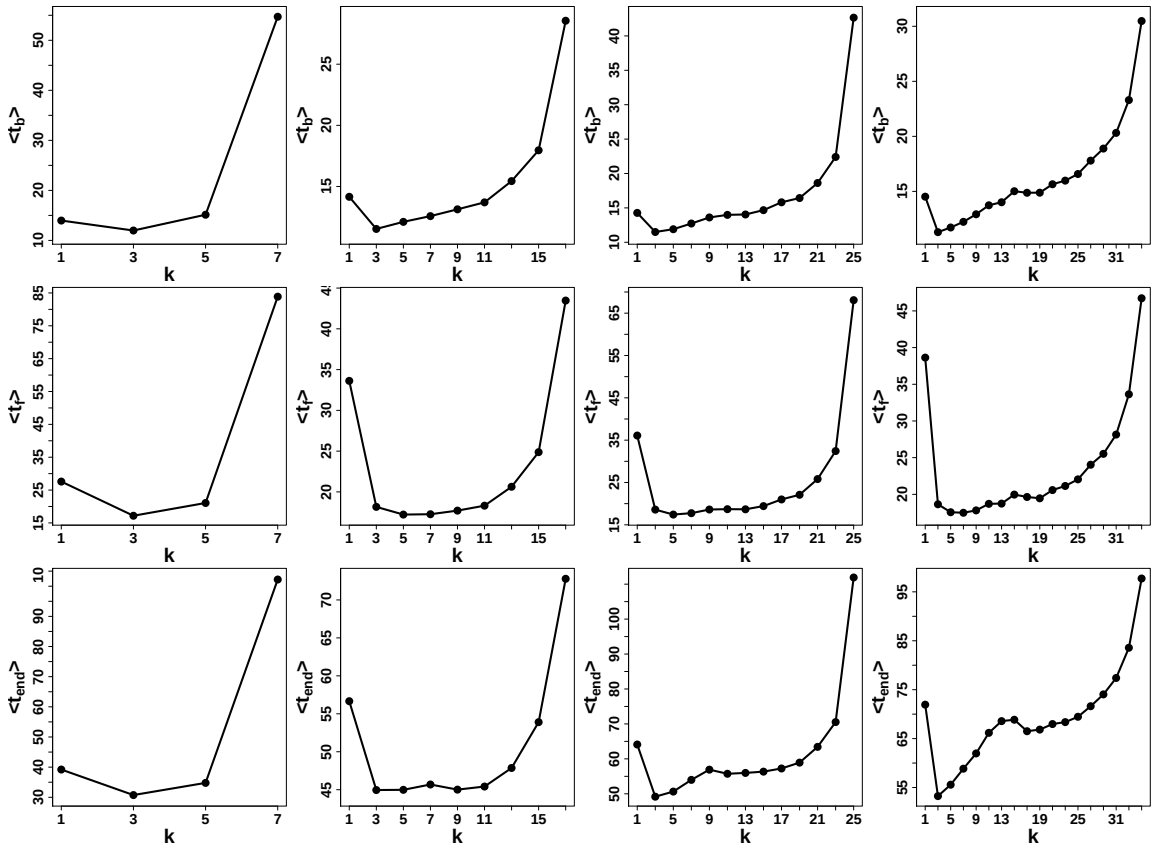

Figure 2.20: Model simulations of the segregation time defined as t_b: the total time an individual spends beeping (first line); t_f: the last time an individual has beeped (second line); t_{end}: the maximum value of t_f for an experiment (third line), against the information parameter k, for groups of 10 (first column), 22 (second column), 30 (third column) and 42 (fourth column) pedestrians.

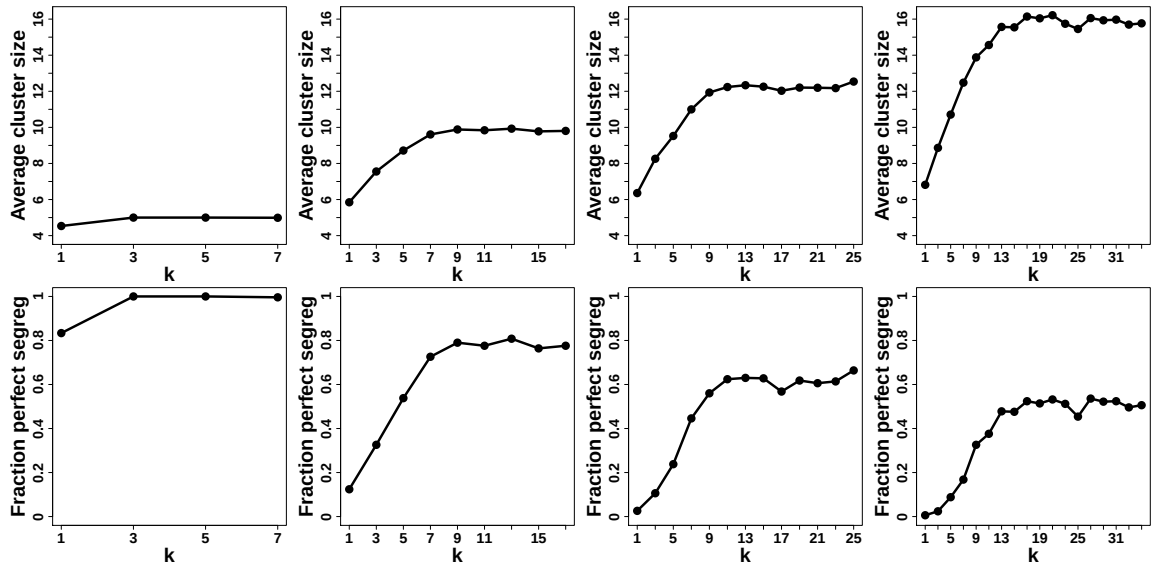

Figure 2.21: Model simulations of the average cluster size (first line) and the fraction of experiments with perfect segregation (second line) against the information parameter k, for groups of 10 (first column), 22 (second column), 30 (third column) and 42 (fourth column) pedestrians. We don't show the other measures again, as they are correlated and don't bring additional insights. Their shape can be deduced easily by looking at Figure 2.19.

- Similarly, the average cluster size at final time doesn't change in proportion to group size. While it equals 5 for $\frac{N}{2} = 5$ (perfect segregation), it is about 10 for $\frac{N}{2} = 11$, about $12 \sim 13$ for $\frac{N}{2} = 15$ and about 16 for $\frac{N}{2} = 21$.
- Correlatively, the fraction of experiments reaching perfect segregation decreases with group size: 100 % for $\frac{N}{2} = 5$, about 80 % for $\frac{N}{2} = 11$, 60 % for $\frac{N}{2} = 15$ and 50 % for $\frac{N}{2} = 21$.

2.5.6 Phase Separation in more Complex Environments

On top of the "*majoritarian*" condition, we tried two other kind of "environments", to look at how more complex tasks would impact the segregation time and quality, and to check if our model (calibrated for the *majoritarian* condition), would reproduce the results obtained in these new conditions.

Since the *exclusive* and *shifted* environments were only done in June 2016, the segregation times observed are about twice those predicted by the model based on the data from September 2015 (as explained in the experimental protocol, section 2.2.2). This kept in mind, Figure 2.22 and 2.23 show that our model is in fair agreement with the data in most cases.

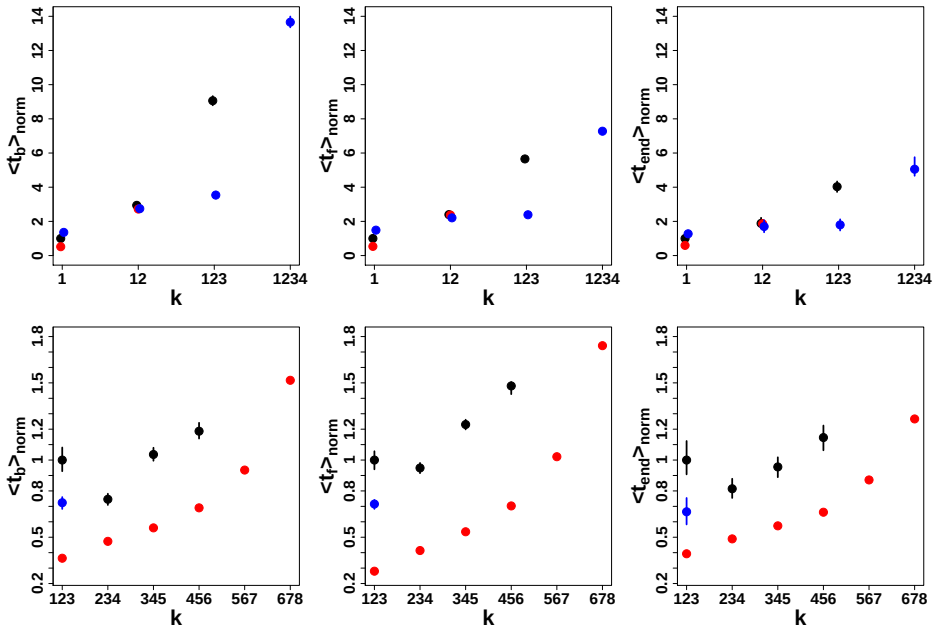

Figure 2.22: Segregation time defined as t_b: the total time an individual spends beeping (first column); t_f: the last time an individual has beeped (second column); t_{end}: the maximum value of t_f for an experiment (third column), against the information parameter k, in the "exclusive" (first line) and "shifted" (second line) conditions (see experimental protocol). All values in each graph are normalized by the value of the corresponding left black dot (which thus serves as a reference). The data are in black and the model simulations are in red. In blue are the experimental data for the condition when subjects were asked to segregate in two clusters. Notice that no red dot appears for the cases '123' and '1234' in the exclusive environment, because the segregation times predicted by the model are extremely long.

As expected, more complex tasks entail longer segregation times. However, our model predicts much longer segregation times for the case '123' in the *exclusive* condition, suggesting the existence of a still not unveiled mechanism. The segregation time was also affected, but less dramatically, when subjects were explicitly asked to segregate in two distinct clusters. Interestingly, this additional condition helped individuals segregate faster in all cases except for $k = 1$ in the exclusive condition (which is the same as $k = 1$ in the majoritarian case).

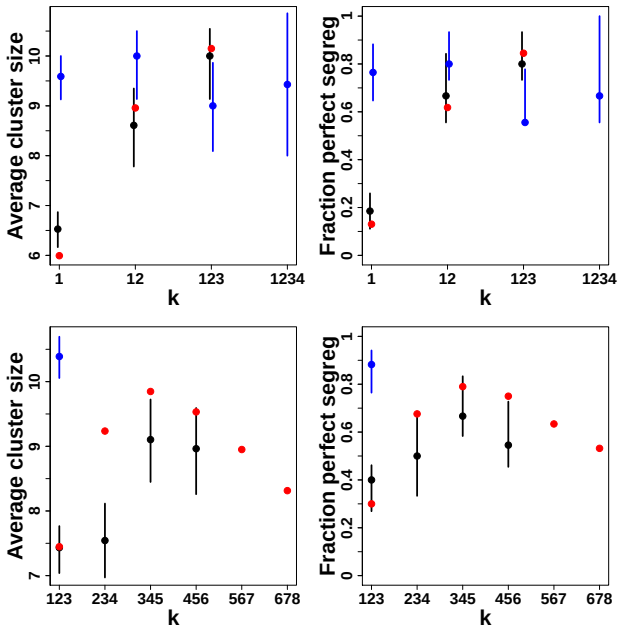

Figure 2.23: Average cluster size (left) and fraction of experiments with perfect segregation (right) against the information parameter k, in the "exclusive" (top) and "shifted" (bottom) conditions (see experimental protocol). The data are in black and the model simulations are in red. In blue are the experimental data for the condition when subjects were asked to segregate in two clusters.

As mentioned above, the clustering process was not affected by the change in segregation time, and our model predictions reproduce the data well. Interestingly, the segregation quality improves with the number of neighbors considered in the *exclusive* criterion. Indeed, the segregation process takes more time, but the average cluster size at final time increases, as does the fraction of experiments reaching perfect segregation.

We also observe a very interesting and non-intuitive pattern in the *shifted* criterion: the segregation quality improves if one considers not the three closest neighbors, but the second, third and fourth closest neighbors, and even better with the third, fourth and fifth closest neighbors. Beyond that, the segregation quality starts decreasing, but slowly and still remaining better than the classic '123' case.

2.6 Conclusion

In this chapter we looked at how collective pedestrian behavior in segregation tasks was affected by varying amounts of information processed by an information-filtering system. We designed an artificial sensory device able to convert complex information in input (colors and positions of all individuals in a closed arena) into a simple bit of information in output (beep or no beep), under specific rules (*majoritarian*, *exclusive* and *shifted* conditions).

The tasks consisted in gathering with pedestrians of the same color. To complete these tasks, subjects only had access to acoustic cues (beeps, output of the information-filtering system) coming from electronic tags attached to their shoulders. They knew neither their own color nor that of the others participants, and were only asked to try not to beep. We defined the quality of a segregation process as better the lesser the segregation time and number of clusters at final time (by analogy with phase separation). Thus, the segregation was considered "perfect" if two distinct clusters were formed at the end of the process (the "phases" were fully separated).

We showed that even though subjects all received, in all conditions tested, the same kind of information (a beep or no beep), they were unconsciously sensitive to subtle variations in upstream information processed by the system. Thus, the segregation time had a U-shaped dependence on k, with an optimal value at $k = 7 \sim 9$ for 22 pedestrians in the *majoritarian* condition. The number of clusters decreased, and the fraction of perfect segregations increased with k, and both reached plateaus at $k = 7 \sim 9$ too in the same condition.

The analysis of the *exclusive* and *shifted* environments interestingly showed that with increased task "complexity", the segregation time worsened (*i.e.* increased), but the segregation quality improved. In particular and quite unexpectedly, when using information from three neighbors (*shifted* environment), the ones to consider in order to optimize the clustering process are the third, fourth and fifth closest.

The same system (tags + sensors + central server) allowed us to geolocalize pedestrians' positions in real-time and analyse their movements with high precision. Using a bottom-up approach (experiments with 1, 2 and 22 pedestrians) and an error minimization based method, we were able to reconstruct the precise functional forms of the *social forces* at play in pedestrian motion.

The model was then used to predict how our results would be affected by changes in group size (groups of 10, 22, 30 and 42 individuals). We found that the optimum value of the segregation time doesn't change with group size, while the plateau and optimum value of k for the segregation quality both increase with group size, in a non-linear fashion. The fraction of perfect segregations also decreases with group size. The model, calibrated for the *majoritarian* criterion, was able to reproduce the results obtained for the two other criteria, underlining its generality.

It is worth emphasizing that subjects had very little comprehensive knowledge of the tasks, and nonetheless managed to complete them with 100 % success. When subjects were explicitly asked to segregate in two groups, the efficiency of the clustering process improved dramatically, especially for low values of k, indicating that indeed, subjects were not aware of the purpose of the tasks (phase separation) in the "classic segregation", which strengthens our results.

This project is not yet entirely complete. In particular, the mechanisms underlying the dependence of segregation quality on group size and the patterns observed in the *exclusive* and *shifted* environments are still not fully understood. Also, the interaction force between two pedestrians is still under investigation.

However, our results already demonstrate the possibility to design information-filtering systems assisting human decision-making, and underline that such systems can be used to *nudge* the collective outcomes of human interactions in desired directions. Of course, they also open the door to manipulation, such that rigorous control and total transparency would be required before they could be applied to real situations.

General Discussion

3.1 Summary of Main Results

In this thesis, we have been interested in understanding the relationship between various amounts of information socially exchanged among individuals in groups and the quality of individual and collective decision making. To that end, we run several experiments in which subjects were faced with two very different kinds of tasks: estimate quantities or segregate in clusters of the same color. In both cases, we rigorously controlled the information that was delivered to the subjects, and looked at the resulting impact on group performance, precisely defined for each task.

3.1.1 Estimation Accuracy in Human Groups

In Chapter 1, subjects estimated various kinds of quantities, and could revise their opinion after having received information from other individuals (the arithmetic or geometric mean of the τ previous estimates). Through a system of artificial agents (see Figure 1.2) inserted into the sequence of estimates (unbeknownst to the subjects) and providing a value of our choice – we called them "informers" or "experts" when the value provided was the true answer – we were able to precisely control the quantity and quality of information shared among individuals in groups, and we studied the resulting impact on individual and collective estimation accuracy.

Distributions of Estimates: Aggregation Methods and Prior Information

We analysed and compared (see section 1.3) distributions of estimates in France (two experiments) and Japan (two experiments) for almost 50 different questions overall, and found that distributions of estimates were closer to Laplace distributions than to Gaussian distibutions, contrary to what was generally found in the literature [60, 117]. Previous works have suggested that distributions of estimates are Generalized Normal Distributions [129], thus uniting Gaussian and Laplace distributions in a common framework, but didn't explain why some distributions were closer to Laplace distributions while others were closer to Gaussian distributions. We provided such an explanation, arguing that the *shape of the distribution* of estimates for a certain quantity is related to the *degree of prior information* of subjects regarding this quantity: the closer a quantity to common intuition (ages, dates, number of objects in an image), the closer the corresponding distribution to a Gaussian distribution, and the farther a quantity to common knowledge (astronomical, physical, biological facts), the closer the corresponding distribution to a Laplace distribution.

We provided arguments for the best aggregation and normalization methods according to the quantity at hand (and the corresponding distribution found). In particular, when a distribution of estimates is close to a Laplace distribution, the Wisdom of Crowds indicator and diversity measure should be respectively the median and the average absolute deviation from the median, rather than the mean and standard deviation (which should be used for Gaussian distributions). These considerations lead us to propose two measures of group performance based of the median rather than the mean, that we called *collective performance* and *collective accuracy*.

Sensitivity to Social Influence: Behavioural, Cognitive and Cultural Differences

We found that subjects' sensitivities to social influence S are very diverse, although one can distinguish five typical behaviours in the distribution of S (see Figure 1.8): keeping one's opinion ($S = 0$), adopting the opinion of the group ($S = 1$), make a compromise ($0 < S < 1$), contradicting the opinion of the group ($S < 0$) or overreacting to it ($S > 1$).

We found consistencies in individuals' answers across questions, which revealed robust personality traits: some subjects – coined "confident" – had a strong tendency to keep their opinions across questions, while others – coined "followers" – tended to trust social information more.

We found that S depends on the distance between personal estimate X_p and social information M (see Figure 1.11): subjects gave more weight to social information as it was farther from their own opinion, which seems inconsistent with the *confirmation bias* that consists in favouring information that confirms one's preexisting beliefs. However, the consistency is recovered if one uses a different phrasing: social information that "confirms" one's personal estimate (bottom of the cusp in Figure 1.11) increases one's confidence[4] in it, and makes thus one more likely to keep it (S closer to 0). Indeed, Figure 1.12 shows that the probability to keep one's opinion is highest when the distance between personal estimate and social information is close to 0, and decreases in favor of the probability to compromise as the distance increases.

Our result is in contrast with another study by Yaniv et al. in 2004, in which the authors found that "the weighting of the advice (*social information*) decreases systematically with distance [between initial estimate and social information]" [62].

In their study, the quantities to estimate were highly demonstrable (dates of historical events within the last three hundred years), which suggests that different cognitive processes may be at play for high and low demonstrability quantities: for high demonstrability quantities, subjects' initial estimates are close to each other and to the true answer (in their experiment they could not be wrong by more than a factor 1.1, in general much less), such that a piece of information is perceived as doubtful or even absurd if it is too far from a subject's initial estimate; on the contrary, for low demonstrability quantities such as those we used, subjects can hardly assess if a piece of social information is absurd or not, whatever its distance from their personal estimate.

In addition, it should be noted that their conclusion was based on only three "points", namely when the social information was at near, intermediate and far distance from the initial (personal) estimate. Therefore, it is not impossible that a closer look at the near to intermediate distances would have shown the same cusp relationship as we found, the difference between their results and ours lying when social information is very far from the personal estimate. In their case (high demonstrability quantities), social information being too far from personal estimate would seem absurd (hence the decrease they found) whereas in our case (low demonstrability quantities), it would seem plausible (hence the plateau we found).

Non-negligible differences were measured in sensitivity to social influence patterns (Figures 1.8 and 1.11) between France and Japan. In particular, Japanese subjects put significantly more weight on social information than French subjects – the median sensitivity to social influence is about 0.6 in Japan (social information is valued higher than the self) against about 0.34 in France (social information is valued lower than the self) – suggesting cultural differences. Time and monetary limitations forbade us to investigate this issue into more details. In particular, we could not run an entire second experiment in Japan, asking subjects to revise their judgement after having received social information. This would have allowed us to directly compare the

[4]Figure 5.10 in Appendix 3.3 indeed shows that sensitivity to social influence and confidence are strongly correlated.

distributions of S in France and Japan when the same questions were asked. Yet, this was not the primary point of our study, and will be kept for future research (see section 3.2.2).

Impact of Information Quantity

We used the above findings (Laplace distribution of estimates, distribution of sensitivities to social influence S, cusp relationship between $\langle S \rangle$ and the distance between personal estimate and social information) to build and calibrate a model of collective estimation processes in human groups, and used it to predict the impact of the *quantity* and *quality* of the information provided ("informers" or "experts" when the information provided is perfectly accurate) to individuals in a group on their individual and collective accuracy.

We showed that collective performance and accuracy improve as the fraction ρ of "experts" – the *quantity* of information provided – increases. In distribution terms, it means that the center of the distribution comes closer to the true value, and its width narrows. Interestingly, we found that while subjects tended to adopt social information more (on average) when they were least accurate, this behaviour lead them to the best improvement in accuracy after social influence. Even when no information was provided ($\rho = 0\,\%$), they were as accurate as the others after social influence.

There is an interesting point to discuss here (see Figure 1.15A): while the accuracy is heterogeneously distributed (the 5 categories are not as accurate; blue points) before social influence, it becomes homogeneous after social influence (the 5 categories are about as accurate; red points). This suggests that the process of exchanging information within the group resulted in an *information transfer* from the most knowledgeable individuals to the least knowledgeable. The mechanism behind this information transfer is not trivial: because more knowledgeable individuals (the most accurate before social influence) tend to be confident and disregard social information – as explained in the previous paragraph – while less knowledgeable individuals (the least accurate before social influence) tend to consider it more, it follows that accurate estimates are less likely to be affected by lower quality estimates than the other way around.

This effect is stronger when $\rho > 0\,\%$: the least knowledgeable/accurate subjects before social influence, by giving some credit to social information, benefited from the "experts" valuable information (even though they were not aware of their existence), and became the most accurate after social influence. Contrariwise, the most knowledgeable/accurate subjects before social influence, because they were too confident in their answer, lost the opportunity to similarly benefit from additional information "grabbed" in the process. As explained in the discussion of Chapter 1, these "experts" can be understood as an external ("environmental") source of information (*e.g.* the media or influential groups) influencing people's opinions. If this information is reliable, then one expects the judgement of people who receive it, and thereby the social information they share, to be improved on average. This is the kind of situation our experiment mimics with the sequential process. The conclusion at this stage is that trusting others is the best strategy when either no information or high-standard information is available.

Impact of Information Quality: Nudging

Obviously, since perfectly accurate information is a quite optimistic assumption, we studied the impact of incorrect information as well ("informers"). Our model predicted a surprising effect (Figure 1.18), later confirmed by preliminary results from the second experiment in France (Figure 1.19): the best collective performance and accuracy after social influence are reached not when perfectly accurate information is provided, but when the information provided *overestimates the truth*, thus compensating the human tendency to underestimate quantities. This

process is a form of *nudging* [46], in the sense that individuals are smoothly – and unconsciously – "pushed" in a certain direction (toward the true value in the present case).

This results exemplify the ambivalence of nudges. As a form of manipulation, they are a way to make use of people's biases for personal profit, such as companies influencing consumers into buying their products, or political parties influencing voters into electing specific candidates. However, they can as well be used to counterbalance those same biases in order to improve people's decision making, and thus potentially enhance their welfare. Typically, nudging technologies could be used to combat filter bubbles and echo chambers rather than favoring them, thus increasing the diversity of information people have access to.

Quite surprisingly, we found that incorrect information was beneficial to collective performance and accuracy over a very large range (up to several orders of magnitude above the truth), reflecting non-linear components in estimation processes (see section 1.7.1). This non-linearity reveals the complexity of the cognitive (tendency to underestimate quantities, tendency to give more credit to information far from one's opinion) and behavioural (distribution of sensitivities to social influence, relationship between confidence and individual accuracy) mechanisms involved in collective estimation processes. In particular, the heterogeneity in individual prior information and social information use are at the core of the non-trivial patterns observed.

Astonishingly, the above conclusion that trusting others is the best strategy in the absence of information or in the presence of accurate information still holds *as long as the information provided is not highly misleading*: subjects who tended to follow social information (*i.e* the least accurate before social influence) were also the most accurate after social influence, over a large range of values of α (the normalized value of the incorrect information): they outperformed the confident individuals (who disregarded social information) over the whole range of incorrect information that we tried, and outperformed the "average" individuals (who consider moderately social information) up to $\alpha = 2$.

The precise mechanisms underlying this surprising results are still under investigation. However, one can already conclude that unless an individual has reasons to believe that her environment provides *highly* inaccurate information, giving credit to social information is statistically the most reliable strategy to improve individual accuracy.

3.1.2 Human Phase Separation

In Chapter 2, we investigated how pedestrians would handle segregation tasks in response to information processed upstream by an information-filtering system. This system – based on a triplet {tags + sensors + central server} – acted as an artificial sensory device able to convert the positions and colors of all individuals walking in a circular arena into a bit of information: the tags attached to a pedestrian's shoulders could emit a beep (1) or not (0).

These acoustic signals were the only cues individuals could use to complete their tasks, which consisted in segregating into clusters of the same color. A tag would emit a beep under specific predefined conditions:

- *majoritarian* condition: a subject's tag would beep if the majority of her k (k =1, 3, 5, 7, 9, 11, 13) nearest neighbors were of a different color from her.

- *exclusive* condition: a subject's tag would beep if *at least one* of her k (k =1, 2, 3, 4) nearest neighbors were of a different color from her.

- *shifted* condition: a subject's tag would beep if the majority of her k^{th}, $(k+1)^{\text{th}}$ and $(k+2)^{\text{th}}$ (k =1, 2, 3, 4) nearest neighbors were of a different color from her.

By controlling k, we controlled the number of closest neighbors, and hence the quantity of information used to compute the output of the system (bit of information). In the *exclusive* condition, varying k consisted also – and in *shifted* condition, only – in manipulating the task complexity.

Model of Pedestrian Motion

The system also permitted to track and record in real time the positions of pedestrians in groups of 1, 2 or 22. We used the data collected to extract the functional forms of the interaction forces – using a reconstruction procedure used in [107] for fish, and described in Appendix 3.3 – and thus build and calibrate a model of pedestrian motion able to reproduce with high fidelity measures as fine as the autocorrelation of speeds, the (confined) diffusion process and distributions of speeds, positions, nearest neighbors and angles to the wall.

Of course, this procedure is based on assumptions. First, it requires a model for the interactions, before their actual form can be extracted. Also, for instance, we assumed that the speed, distance and angular components of the interaction with the wall were decoupled. Still, we didn't have much more assumptions to make, and the model we chose is very reasonable given the task pedestrians had to accomplish (walking randomly inside a circular arena). This procedure thus allows a high degree of generality.

Impact of the Amount of Information on Phase Separation

To look at the impact of the amount of information used by the information-filtering system (quantified by the parameter k), we defined quantifiers of group performance. We say that a segregation process is better as segregation time is shorter and – by analogy with phase separation – the number of clusters at final time is smaller. According to this definition, the segregation was considered "perfect" if only two clusters remained at the end of the process (the "phases" were fully separated).

We found that the segregation time had a U-shaped dependence on k, with an optimal (minimal) value at $k = 7 \sim 9$ for groups of 22 pedestrians in the majoritarian condition. This suggests that too much information ($k > 9$) makes it more difficult to complete the task, as the information becomes too noisy.

We didn't observe such an optimum value for the clustering quality at final time, but rather a saturation (plateau) at the same value $k = 7 \sim 9$. All measures chosen showed the same saturation at the same value of k. The phase separation was thus very incomplete for low values of k, and improved with k up to $k = 7 \sim 9$, for which pedestrians segregated perfectly in about 80% experiments.

Our model showed that these patterns remain globally stable when the group size changes, and that the optimum value of k for the segregation time doesn't depend on group size. Yet, the clustering process does depend on group size. The fraction of experiments where the segregation is perfect decreases with higher group sizes, and the average cluster size at final time increases with group size, but not in proportion of it. Similarly, the value of k for which the plateau is reached doesn't vary linearly with group size. The precise mechanisms underlying these results are still under investigation.

Impact of Task Complexity on Phase Separation

The *exclusive* and *shifted* conditions allowed us to play with the "complexity" of the tasks. For instance, in the *exclusive* condition, a pedestrian's tag would stop beeping if his k nearest neighbors were *all* of the same color as him. One thus expects the task to be more difficult as k increases (higher probability to beep). Similarly, one expects the *shifted* condition to be less stable as k increases. Indeed, it should be more difficult to form a one-color cluster if one tries to stabilize (stop beeping) according to the positions of neighbors that are not the nearest ones, because the latter (who are in between) may well be of the other color.

The analyses showed that the segregation time increased with task "complexity", as could be expected. Yet, quite surprisingly, the clustering quality improved with k for both conditions, suggesting that introducing noise in the task could actually stabilize the final state, even tough it takes more time to reach it. Similar positive effects of noise on collective performance in human groups, especially in coordination problems, have been observed elsewhere [154].

Interestingly, our model simulations showed that the clustering process was optimized when using information from the third, fourth and fifth closest neighbors in the *shifted* condition. The mechanisms underlying these results are not yet fully understood.

Manipulating the Outcome of Collective Processes

The main focus of this project was not to emphasize how one could use information quantity and quality to improve phase separation in human groups, but rather to show that it is possible to couple pedestrian activity with a device that controlled and filtered the information available before transmitting a "simplified" (condensed) information to pedestrians, helping them perform a phase-separation-like task.

Indeed, phase separation in groups of pedestrians has little or no real life applications. It is rather an artificial situation designed in order to make quantitative measurements for the point we want to make. Correlatively, we proposed ad hoc definitions of group performance relative to this specific type of tasks, and showed how these performances were affected by the information used and processed by the device.

By showing that it is possible to use such systems to steer the outcome of collective human processes in desired directions, this work opens the way to the development of computer based systems aiming at enhancing collective decision-making in human groups in more general situations. In particular, it would be interesting in future research to design such systems adapted to more realistic situations, such as artificial crowd panics, and see if they could help nudging individual behaviors toward safe (or safer) collective outcomes.

In the next section, I will present some research ideas I intend to undertake as a post-doctoral fellow at Max Planck Institute for Human Development, under the supervision of Dr. Ralf Kurvers. These ideas are all related to the first Chapter of this thesis, as it is more in line with the Center for Adaptive Rationality's line of research, as well as with my personal interests.

3.2 Future Research Directions

This research raised a number of questions and was the outset of a broader plan aiming at understanding the cognitive and behavioral processes underlying the emergence of collective intelligence in human groups. Below I briefly present the research directions I intend to follow as a post-doctoral fellow.

3.2.1 Theory of Social Impact

In Chapter 1, our model predicted that collective performance and accuracy should depend on τ (the number of previous estimates whose geometric mean was given as social information), even if subjects didn't know the value of τ (namely the number of other individuals involved in the social information provided).

However, this dependence was too little to be observed experimentally, such that we couldn't draw any conclusion as to the impact of the number of individuals providing social information on group performance. Previous work suggested that social information impacts people as a power function of the number of individuals expressing it [155]. But they didn't look at how a group can use this information to improve its performance collectively.

One future project will therefore be to look at the impact of the number τ of information sources on estimation accuracy in a similar framework (low demonstrability estimation tasks). We will compare a condition in which all τ values are provided to a condition in which only the geometric mean of the τ values is provided (but where individuals would know the value of τ).

I believe the number of individuals providing social information, the confidence with which they express their opinion and their reputation to be among the major determinants of social information use at the individual scale. Therefore, in a longer term, I intend to encompass these three elements at least in a common framework that would be a first step toward a realistic theory of social impact. Such a theory seems of utmost importance to understand opinion formation and decision making processes, particularly online with the recent rise of interactivity (likes/dislikes, rates, comments...) in platforms such as Facebook, Youtube, Netflix, Amazon and countless others.

3.2.2 Determinants of Social Information Use

Figures 1.8, 1.11 and 5.10 suggested important cultural differences between France and Japan, although they were insufficient to ascertain it (because the questions asked in both countries were different). In particular, Japanese subjects had a much higher tendency to trust social information (the median sensitivity to social influence in Japan was 0.6, against 0.35 in France) and correlatively a lower confidence in their answer.

This lead us to hypothesize that sensitivity to social influence may be a good proxy for cultural differences. Indeed, we expect very different societies (in subsistence style, hierarchy structure, market integration, social network structure...) to exhibit different patterns of social information use. Recent work on social learning supported this assumption, showing that the reliance on social information in learning processes varies with populations' subsistence style (*e.g.* pastoralists have a markedly higher propensity for social learning than horticulturalists) [156].

Yet, the precise factors driving these differences are still unknown: if our hypothesis is correct, what lead Japanese subjects to give more weight to social information than French subjects? To try to answer this question, we intend to ask members of indigenous tribes, with very different cultures and habits, to participate to estimation tasks experiments. The objective will be to relate the patterns (distributions) of sensitivities to social influence to some criteria characteristic of these societies (*e.g.* individualistic versus collectivist societies, egalitarian versus hierarchical societies...).

This project will focus on the determinants of social impact at a population scale, thus complementing the project described above.

3.2.3 Assignment of Social Information

In our experiment, subjects were provided as social information the average estimate of τ previous participants, which were not selected. Similarly, in most other works presented in this paper, the social information provided was either one or the average of several randomly chosen participants' estimates.

However, one could argue that this may not be the optimal way for individuals to exchange information in a group. Indeed, results in section 1.7 suggest that it is possible to improve the final estimates of individuals by providing them with some information that overestimates the truth. Would it then be possible to improve the outcome of collective estimation processes by changing the way social information is exchanged?

Also, in our experiments external information was provided to the group ("informers"), thus influencing social information itself and the collective outcome. This is useful to mimic situations in which decision-making is sequential and individuals in a group can gather information from outside in the process. However, not all situations match this scheme, and it would be interesting to find conditions under which decision-makers can significantly improve their individual and collective accuracy after social influence, using only information from their peers.

To that end, we intend to run experiments in which subjects have to complete non-sequential estimation tasks. After providing their personal estimates, they will receive as social information one of the three following sets of estimates:

- τ random estimates taken from other individuals in the group;
- the τ estimates that are closest to the median estimate of the group (more precisely the τ estimates which logarithm is closest to the median log-estimate). This condition is justified by the fact that the median (log-)estimate is statistically better than individual (log-)estimates (Wisdom of Crowds);
- the τ estimates that are (which logarithm is) closest to *an overestimated value of* the median (log-)estimate, which takes into account the fact that the median estimate of a group is generally negatively biased (tendency to underestimate quantities; see Figure 1.3). We thus expect to improve further individual and collective accuracy by partly compensating this bias.

At the time of writing these lines, the project has already started. All mathematical details and justifications will be given in a publication to come. Note that we didn't use any reference to the true value of quantities to estimate, nor any information coming from outside the group ("informers"). Individuals have no other information than their own distributed information to complete the task.

3.2.4 Estimates, Opinions and Decision-Making

A legitimate concern regarding estimation tasks is to what extent they represent or even mimic real life situations, such that one can question the generality of the results found in Chapter 1:

1. Most real-life decision-making situations imply to choose among a discrete, limited, and often small set of alternatives, at variance with the continuous and potentially infinite number of possible estimates in our tasks. Often, the alternatives are not estimates, but options, which are often not even numerical (new policies, release of a new product, choice of a new doctor, of a movie...).

In such contexts, models for continuous choices don't apply as such. In particular, aggregation measures such as the mean or median are not or ill defined for discrete choice problems, especially if the number of choices is small. For example, in a situation where two choices are available and are 1 and 3, the average and median would be 2, which is not even one of the possible choices. Therefore, alternatives are to be found, among which voting seems a viable candidate [56]. It has been suggested that (majority) voting is the discrete choices analogous to averaging for continuous choices [157].

It would be an interesting project to find a way to adapt the model presented in Chapter 1 to discrete choices problems, and then try it against discrete choice experimental tasks.

2. A lot of decision-making processes don't involve a "true" answer, such that choices are *opinions* rather than *estimates*. For such "truthless" problems, notions such as "group performance" make little sense, but social influence processes are at play, and constitute a common ground for Wisdom of Crowds (estimates-based) and opinion dynamics research areas.

   Despite this shared interest, they differ in their focus and methods: the former intents to unveil the mechanisms underlying collective behaviour (social influence processes in particular), while the latter focuses on the time evolution of these behaviours (social phenomena such as polarization of opinions, emergence of consensus, spread of minority opinions... see example surveys in [158, 159]). The former's results are experimentally grounded while the latter relies mostly on simulations.

   Both approaches are quite complementary, but don't "communicate" enough in the sense that the assumptions (*e.g.* bounded confidence) used in the main opinion dynamics models (Degroot [160], Axelrod [161], Hegselmann-Krause [158] or Deffuant-Weisbuch [162] models) should be directly inspired (if not taken) from the findings of estimation-based research, which is not always the case.

   It would therefore be interesting to try to add a time component to the model proposed in Chapter 1 and compare the results to those obtained in the reference papers on opinion dynamics.

3.3 Concluding Words

In this PhD we undertook two major projects aiming to investigate the impact of information on decision making in human groups, in two very different setups: one based on estimations tasks, and the other on pedestrian motion. These two main projects will lead to four papers at least, one of which has already been published in PNAS in November 2017 [163], and the others will be submitted soon.

This PhD was an incredible opportunity for me to develop a large variety of competencies and increase my knowledge in various domains. It was not always easy, but overall it was a unique and rewarding experience that I would recommend to anyone interested in pursuing his studies beyond Master degree.

Thanks to my supervisors and other collaborators, I believe I acquired the necessary skills to start a career as a researcher, and I definitely intend to continue in this direction. My main interests undoubtedly go to computational social sciences, and in particular everything related to cognition and decision making. I think in this respect at least, the Center for Adaptive Rationality is a perfect place to continue my research activities.

Bibliography

[1] Herbert A. Simon. *Models of Bounded Rationality*. MIT Press, 1982.

[2] Daniel Kahneman. *Thinking: Fast and Slow*. New York: Farrar, Straus and Giroux, 2011.

[3] Dan Ariely. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. HarperCollins, 2008.

[4] Allen G. Schick and Lawrence A. Gordon. Information overload: A temporal approach. *Accounting, Organization and Society*, 15(3):199–220, 1990.

[5] Torkel Klingberg. *The Overflowing Brain: Information Overload and the Limits of Working Memory*. 2008.

[6] Juho Salminen. Collective intelligence in humans: A literature review. *arXiv:1204.3401*, 2012.

[7] Eric Bonabeau. Decisions 2.0: the power of collective intelligence. *MIT Sloan Management Review, Cambridge*, 50(2):45–52, 2009.

[8] Anita Williams Woolley, Ishani Aggarwal, and Thomas W. Malone. Collective intelligence and group performance. *Current Directions in Psychological Science*, 24(6):420–424, 2015.

[9] James Surowiecki. *The Wisdom of Crowds*. Anchor Books, 2005.

[10] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm intelligence: from natural to artificial systems*. New York, NY: Oxford University Press, Santa Fe Institute Studies in the Sciences of Complexity, 1999.

[11] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.

[12] Milton Friedman and Leonard J. Savage. The utility analysis of choices involving risk. *Journal of Political Economy*, 56(4):279 – 304, 1948.

[13] Kenneth J. Arrow. *Essays in the theory of risk-bearing*. Chicago: Markham, 1971.

[14] Herbert A. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138, 1956.

[15] Daniel G. Goldstein and Gerd Gigerenzer. Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109(1):75–90, 2002.

[16] Vernon L. Smith. Constructivist and ecological rationality in economics. *The American Economic Review*, 93(3):465–508, 2003.

[17] Martie Haselton et al. Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, 27(5):733–763, 2009.

[18] Ralph Hertwig and Stefan M. Herzog. Fast and frugal heuristics: Tools of social rationality. *Social Cognition*, 27(5):661–698, 2009.

[19] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1974.

[20] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

[21] Ward Edwards. *Conservatism in human information processing.* Cambridge University Press, 1982.

[22] Massimo Piattelli-Palmarini. *Inevitable illusions. How mistakes of reason rule our minds.* New York: Wiley, 1994.

[23] Gerd Gigerenzer and Reinhard Selten, editors. *Bounded Rationality: The Adaptive Toolbox.* Dahlem Workshop Reports, 2001.

[24] Gerd Gigerenzer and Henry Brighton. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1:107–143, 2009.

[25] Gerd Gigerenzer and Daniel G. Goldstein. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4):650–669, 1996.

[26] Joyce Ehrlinger, Wilson O. Readinger, and Bora Kim. Decision-making and cognitive biases. *Encyclopedia of Mental Health*, 2016.

[27] Solomon E. Asch. Opinions and social pressure. *Scientific American*, 193(5):31–35, 1955.

[28] Stanley Milgram. Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67(4):371–378, 1963.

[29] Stanley Milgram. The small world problem. *Journal of Economic Behavior and Organization*, 1(1):61–67, 1967.

[30] Stanley Milgram, Leonard Bickman, and Lawrence Berkowit. Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology*, 13(2):79–82, 1969.

[31] Jacques Gravereau et Jacques Trauman, editor. *Crises financières.* Economica, 2001.

[32] André Orléan. Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion. *Journal of Economic Behavior and Organization*, 28:257–274, 1995.

[33] Thomas D. Wilson. Models in information behaviour research. *The Journal of Documentation*, 55(3):249–270, 1999.

[34] Natasha A. Karlova and Karen E. Fisher. A social diffusion model of misinformation and disinformation for understanding human information behaviour. *Information Research*, 18(1):paper 573, 2013.

[35] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–236, 2017.

[36] Stephan Lewandowsky, Ullrich K.H. Ecker, and John Cook. Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, 6:353–369, 2017.

[37] Michela Del Vicario et al. The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, 113(3):554–559, 2016.

[38] Vahed Qazvinian et al. Rumor has it: Identifying misinformation in microblogs. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1589–1599, 2011.

[39] Aditi Gupta et al. Tweetcred: Real-time credibility assessment of content on twitter. *Social Informatics. SocInfo 2014. Lecture Notes in Computer Science*, 8851:228–243, 2014.

[40] Jacob Ratkiewicz et al. Detecting and tracking political abuse in social media. *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*, pages 297–304, 2011.

[41] Xin Luna Dong et al. Knowledge-based trust: estimating the trustworthiness of web sources. *Proceedings of the VLDB Endowment*, 8(9):938 – 949, 2015.

[42] Elizabeth F. Loftus and Hunter G. Hoffman. Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General*, 118(1):100–104, 1989.

[43] Steven J. Frenda, Rebecca M. Nichols, and Elizabeth F. Loftus. Current numbers and advances in misinformation research. *Current Directions in Psychological Science*, 20(1):20–23, 2011.

[44] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.

[45] Christopher J. Fox. *Information and misinformation: An investigation of the notions of information, misinformation, informing, and misinforming.* 1983.

[46] Richard Thaler and Cass Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness.* 2008.

[47] Francis Galton. Vox populi. *Nature*, 75:450–451, 1907.

[48] Francis Galton. One vote one value. *Nature*, 75:414, 1907.

[49] J. Scott Armstrong. *Principles of forecasting: a handbook for researchers and practitioners.* Norwell, MA: Kluwer Academic Publishers, 2001.

[50] Karsten Hueffer et al. The wisdom of crowds: Predicting a weather and climate-related event. *Judgment and Decision Making*, 8(2):91–105, 2013.

[51] Kenneth J. Arrow et al. The promise of prediction markets. *Science*, 320(5878):877–878, 2008.

[52] Justin Wolfers and Eric Zitzewitz. Prediction markets. *Journal of Economic Perspectives*, 18(2):107–126, 2004.

[53] Justin Wolfers and Eric Zitzewitz. Interpreting prediction market prices as probabilities. *NBER Working Paper Series*, Working Paper 12200, 2006.

[54] Max Wolf et al. Collective intelligence meets medical decision-making: The collective outperforms the best radiologist. *Plos One*, 10(8):e0134269, 2015.

[55] Ralf H.J.M. Kurvers et al. Boosting medical diagnostics by pooling independent judgments. *Proceedings of The National Academy of Sciences of the United States of America*, 113(31):8777 – 8782, 2016.

[56] Marquis de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. 1785.

[57] David Austen-Smith and Jeffrey S. Banks. Information aggregation, rationality, and the condorcet jury theorem. *The American Political Science Review*, 90(1):34–45, Mar 1996.

[58] Jan Lorenz et al. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22):9020–9025, 2011.

[59] Mehdi Moussaïd et al. Social influence and the collective dynamics of opinion formation. *PLoS ONE*, 8(11):e78433, 2013.

[60] Gabriel Madirolas and Gonzalo G. de Polavieja. Improving Collective Estimations Using Resistance to Social Influence. *PLOS Computational Biology*, 11(11):e1004594, 2015.

[61] Effects of social influence on the wisdom of crowds. *Proceedings, CI 2012*.

[62] Ilan Yaniv. Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, 93(1):1–13, 2004.

[63] Andrew J. King et al. Is the true 'wisdom of the crowd' to copy successful individuals? *Biology Letters*, 8(2):197–200, 2012.

[64] Nadav Klein and Nicholas Epley. Group discussion improves lie detection. *Proceedings of the National Academy of Sciences of the United States of America*, 112(24):7460–7465, 2015.

[65] Mehdi Moussaïd et al. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PLoS ONE*, 5(4):1–7, 2010.

[66] Mehdi Moussaid et al. Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in cognitive science*, 1(3):469 – 497, 2009.

[67] Audrey Dussutour et al. Optimal traffic organization in ants under crowded conditions. *Nature*, 428:70–73, 2004.

[68] Nigel R. Franks and Charles R. Fletcher. Spatial patterns in army ant foraging and migration: Eciton burchelli on barro colorado island. *Behavioral Ecology and Sociobiology*, 12(4):261–270, 1983.

[69] Andrea Perna et al. The topological fortress of termites. *Bio-Inspired Computing and Communication. Lecture Notes in Computer Science*, 5151:165–173, 2008.

[70] Simon Garnier, Jacques Gautrais, and Guy Theraulaz. The biological principles of swarm intelligence. *Swarm Intelligence*, 1(1):3 – 31, 2007.

[71] Tatsuya Kameda et al. Is consensus-seeking unique to humans? a selective review of animal group decision-making and its implications for (human) social psychology. *Group Processes & Intergroup Relations*, 15(5):673–689, 2012.

[72] John H. Miller and Scott E Page. *Complex adaptive systems: an introduction to computational models of social life*. Princeton, N.J. : Princeton University Press, 2007.

[73] Eric Bonabeau et al. Self-organization in social insects. *Trends in Ecology and Evolution*, 12(5):188–193, 1997.

[74] Scott Camazine et al. *Self-Organization in Biological Systems*. 2001.

[75] Ralph Beckers et al. Collective decision making through food recruitment. *Insectes sociaux, Paris*, 37(3):258–267, 1990.

[76] Ralph Beckers, Jean-Louis Deneubourg, and Simon Goss. Modulation of trail laying in the ant lasius niger (hymenoptera: Formicidae) and its role in the collective selection of a food source. *Journal of Insect Behavior*, 6(6):751–759, 1993.

[77] Simon Goss et al. Self-organized shortcuts in the argentine ant. *Naturwissenschaften*, 76:579–581, 1989.

[78] Guy Theraulaz and Eric Bonabeau. A brief history of stigmergy. *Artificial life*, 5(2):97–116, 1999.

[79] Marco Dorigo, Eric Bonabeau, and Guy Theraulaz. Ant algorithms and stigmergy. *Future Generation Computer Systems*, 16:851–871, 2000.

[80] Jennifer R. Riley et al. The flight paths of honeybees recruited by the waggle dance. *Nature*, 435:205–207, 2005.

[81] Jacobus C. Biesmeijer and Thomas D. Seeley. The use of waggle dance information by honey bees throughout their foraging care. *Behavioral Ecology and Sociobiology*, 59(1):133–142, 2005.

[82] J. Stephen Lansing. Complex adaptive systems. *Annual Review of Anthropology*, 32:183–204, 2003.

[83] Eric Bonabeau. Marginally stable swarms are flexible and efficient. *Journal de Physique I*, 6(2):309–324, 1996.

[84] Alfred J. Lotka. Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Science of the United States of America*, 6(7):410–415, 1920.

[85] Roger Arditi and Lev R. Ginzburg. Coupling in predator-prey dynamics: Ratio-dependence. *Journal of Theoretical Biology*, 139(3):311–326, 1989.

[86] Alan A. Berryman. The origins and evolution of predator-prey theory. *Ecology*, 73(5):1530–1535, 1992.

[87] Iain D. Couzin et al. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1):1 – 11, 2002.

[88] Daniel S. Calovi et al. Swarming, schooling, milling: phase diagram of a data-driven fish school model. *New Journal of Physics*, 16:015026, 2014.

[89] Benjamin Pettit et al. Interaction rules underlying group decisions in homing pigeons. *Journal of the Royal Society Interface*, 10(89):20130529, 2013.

[90] Andrea Procaccini et al. Propagating waves in starling, sturnus vulgaris, flocks under predation. *Animal Behaviour*, 82(4):759 – 765, 2011.

[91] Sylvain Toulet et al. Imitation combined with a characteristic stimulus duration results in robust collective decision-making. *Plos One*, 10(10):e0140188, 2015.

[92] Francesco Ginelli et al. Intermittent collective dynamics emerge from conflicting imperatives in sheep herds. *Proceedings of the National Academy of Sciences of the United States of America*, 112(41):12729–12734, 2015.

[93] Dirk Helbing, Joachim Keltsch, and Péter Molnár. Modelling the evolution of human trail systems. *Nature*, 388:47–50, 1997.

[94] Dirk Helbing et al. Self-organizing pedestrian movement. *Environment and Planning B: Planning and Design*, 28:361–383, 2001.

[95] Dirk Helbing, Anders Johansson, and Habib Zein Al-Abideen. Dynamics of crowd disasters: An empirical study. *Physical Review E*, 75:046109, 2007.

[96] Dirk Helbing, Illés, and Tamás Vicsek. Simulating dynamical features of escape panic. *Nature*, 407:487–490, 2000.

[97] Andreas Schadschneider et al. Evacuation dynamics: Empirical results, modeling and applications. *Encyclopedia of Complexity and System Science*, pages 3142–3176, 2009.

[98] Franck Schweitzer. *Brownian Agents and Active Particles. Collective Dynamics in the Natural and Social Sciences.* Berlin: Springer (Springer Series in Synergetics), 2003.

[99] Dirk Helbing and Péter Molnár. Social force model for pedestrian dynamics. *Physical Review E*, 51(5):4281 – 4286, 1995.

[100] Dirk Helbing. *Quantitative Sociodynamics : Stochastic Methods and Models of Social Interaction Processes.* Springer, 2nd edition., 1995.

[101] Dirk Helbing et al. Active walker model for the formation of human and animal trail systems. *Physical Review E*, 56(3):2527–2539, 1997.

[102] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences of the United States of America*, 108(17):6884 – 6888, 2011.

[103] Mehdi Moussaid et al. Experimental study of the behavioural mechanisms underlying self-organization in human crowds. *Proceedings of the Royal Society B*, 276:2755–2762, 2009.

[104] Serge P. Hoogendoorn and Winnie Daamen. Pedestrian behavior at bottlenecks. *Transportation Science*, 39(2):147–159, 2005.

[105] Gianluca Antonini, Michel Bierlaire, and Mats Weber. Discrete choice models of pedestrian walking behavior. *Transportation Research Part B: Methodological*, 40(8):667–687, 2006.

[106] Michele Ballerini et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences of the USA*, 105(4):1232–1237, 2008.

[107] Daniel S. Calovi et al. Disentangling and modeling interactions in fish with burst-and-coast swimming reveal distinct alignment and attraction behaviors. *PLoS Computational Biology*, 14(1):e1005933, 2018.

[108] Chrysanthos Dellarocas. The digitization of word-of-mouth: Promise and challenges of online feedback mechanisms. *Management Science*, 49(10):1407–1424, 2003.

[109] Meeyoung Cha et al. Measuring user influence in twitter: The million follower fallacy. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 10–17, 2010.

[110] Bernard J. Jansen et al. Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, (60):2169–2188, 2009.

[111] Bruno Gonçalves and Nicola Perra. *Social phenomena: From data analysis to models*. Heidelberg, New-York: Springer International Publishing AG., 2015.

[112] Robert M. Bond et al. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489:295–298, 2012.

[113] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026, 1992.

[114] Abhijit V. Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, 107:797–817, 1992.

[115] Patrick R. Laughlin. *Group problem solving*. Princeton University Press, 2011.

[116] Tatsuya Kameda, R. Scott Tindale, and James H. Davis. *Cognitions, preferences, and social sharedness: Past, present, and future directions in group decision making*. Schneider & J. Shanteau (Eds.), Emerging perspectives on judgment and decision research. Cambridge: Cambridge University Press, 2009.

[117] Pavlin Mavrodiev, Claudio J. Tessone, and Frank Schweitzer. Quantifying the effects of social influence. *Scientific Reports*, 3:1360, 2013.

[118] Corentin V. Kerckhove et al. Modelling influence and opinion evolution in online collective behaviour. *PLoS ONE*, 11(6):e0157685, 2016.

[119] Yu Luo, Garud Iyengar, and Venkat Venkatasubramanian. Social influence makes self-interested crowds smarter: an optimal control perspective. *IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS*, 5(1):200 – 209, 2018.

[120] Jolyon J. Faria et al. Leadership and social information use in human crowds. *Animal Behaviour*, 79(4):895–901, 2010.

[121] Stanislas Dehaene et al. Log or linear? distinct intuitions of the number scale in western and amazonian indigene cultures. *Science*, 320(5880):1217–1220, 2008.

[122] Clare Harries, Ilan Yaniv, and Nigel Harvey. Combining advice: the weight of a dissenting opinion in the consensus. *Journal of Behavioral Decision Making*, 17(5):333–348, 2004.

[123] Andrés Chacoma and Damián H. Zanette. Opinion formation by social influence: From experiments to modeling. *PLoS One*, 10(10):e0140406, 2015.

[124] Tarow Indow and Masashi Ida. Scaling of dot numerosity. *Perception & Psychophysics*, 22(3):265–276, 1977.

[125] Pierre-Simon Laplace. Mémoire sur la probabilité des causes par les évènements. *Mémoires de l'Academie Royale des Sciences Presentés par Divers Savan*, 6:621–656, 1774.

[126] Edwin B. Wilson. First and second laws of error. *Journal of the American Statistical Association*, 18(143):841–851, 1923.

[127] Paul R. Rider. A generalized law of error. *Journal of the American Statistical Association*, 19(146):217–220, 1924.

[128] Saralees Nadarajah. A generalized normal distribution. *Journal of Applied statistics*, 32(7):685–694, 2005.

[129] Miguel Sousa Lobo and Dai Yao. Human judgment is heavy tailed: Empirical evidence and implications for the aggregation of estimates and forecasts. *INSEAD working paper series*, 2010.

[130] Johannes H. C. Lisman and Martien C. A. van Zuijlen. Note on the generation of most probable frequency distributions. *Statistica Neerlandica*, 26(1):19–23, 1972.

[131] John M. Keynes. The principal averages and the laws of error which lead to them. *Journal of the Royal Statistical Society*, 74(3):322–331, 1911.

[132] Jack B. Soll and Richard P. Larrick. Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3):780–805, 2009.

[133] Philip Ball. *Why society is a complex matter: meeting twenty-first century challenges with a new kind of science*. Springer, Berlin, 2012.

[134] Irving Lester Janis. *Groupthink: Psychological studies of policy decisions and fiascoes*. Houghton Mifflin Boston, 1982.

[135] Dirk Helbing. Globally networked risks and how to respond. *Nature*, 497(7447):51–59, May 2013.

[136] Fatos Xhafa and Nik Bessis, editors. *Inter-cooperative Collective Intelligence: Techniques and Applications*. Studies in Computational Intelligence. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[137] David J.T. Sumpter and Madeleine Beekman. From nonlinearity to optimality: pheromone trail foraging by ants. *Animal Behavior*, 66(2):273 – 280, 2003.

[138] Daniel R. MacNulty et al. Influence of group size on the success of wolves hunting bison. *Plos One*, 9(11):e112884, 2014.

[139] Young-Ho Eom et al. Network-based model of the growth of termite nests. *Physical Review E*, 92(6):062810, 2015.

[140] Stefano Marras and Paolo Domenici. Schooling fish under attack are not all equal: Some lead, others follow. *Plos One*, 8(6):e65784, 2013.

[141] Anne E. Magurran. The adaptive significance of schooling as an anti-predator defence in fish. *Annales Zoologici Fennici*, 27(2):51 – 56, 1990.

[142] Robert Mach and Frank Schweitzer. Modeling vortex swarming in daphnia. *Bulletin of Mathematical Biology*, 69:539 – 562, 2007.

[143] Dirk Helbing. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, 73:1067 – 1141, 2001.

[144] Pablo Jensen. A network-based prediction of retail stores commercial categories and optimal locations. *Physical Review E*, 74:035101, 2006.

[145] Lee Jooyong et al. Modeling lane formation in pedestrian counter flow and its effect on capacity. *KSCE Journal of Civil Engineering*, 20(3):1099–1108, 2016.

[146] Yi zhou Tao and Li yun Dong. Investigation on lane-formation in pedestrian flow with a new cellular automaton model. *Journal of Hydrodynamics, Series B*, 28(5):794–800, 2016.

[147] Wei Guoa, Xiaolu Wanga, and Xiaoping Zheng. Lane formation in pedestrian counterflows driven by a potential field considering following and avoidance behaviours. *Physica A: Statistical Mechanics and its Applications*, 432:87–101, Aug 2015.

[148] Dirk Helbing et al. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation Science*, 39(1):1–24, 2005.

[149] Nirajan Shiwakoti et al. Understanding crowd panic at turning and intersection through model organisms. *Pedestrian and Evacuation Dynamics 2012*, pages 1175–1183, 2014.

[150] Dirk Helbing et al. Simulation of pedestrian crowds in normal and evacuation situations. *Pedestrian and Evacuation Dynamics*, pages 21–58, 2002.

[151] Serge P. Hoogendoorn and Winnie Daamen. Microscopic calibration and validation of pedestrian models: Cross-comparison of models using experimental data. *Traffic and Granular Flow'05*, pages 329–340, 2007.

[152] Anders Johansson, Dirk Helbing, and Pradyumn K. Shukla. Specification of the social force pedestrian model by evolutionary adjust- ment to video tracking data. *Advances in Complex Systems*, 10:271–288, 2007.

[153] Jacques Gautrais et al. Deciphering interactions in moving animal groups. *PLoS Computational Biology*, 8(9):e1002678, 2012.

[154] Hirokazu Shirado and Nicholas A. Christakis. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature*, 545:370 – 374, 2017.

[155] Bibb Latané. The psychology of social impact. *American psychologist*, 36(4):343–356, 1981.

[156] Luke Glowacki and Lucas Molleman. Subsistence styles shape human social learning strategies. *Nature Human Behaviour*, 1:0098, 2017.

[157] Andres Laan, Gabriel Madirolas, and Gonzalo G. de Polavieja. Rescuing collective wisdom when the average group opinion is wrong. *Frontiers in Robotics and AI*, 4:56, 2017.

[158] Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence, models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5:3, 2002.

[159] Jan Lorenz. Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C*, 18(12):1819–1838, 2007.

[160] Morris H. Degroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[161] Robert Axelrod. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, 41(2):203–226, 1997.

[162] Gérard Weisbuch et al. Interacting agents and continuous opinions dynamics. *Heterogenous Agents, Interactions and Economic Performance. Lecture Notes in Economics and Mathematical Systems*, 521:225–242, 2003.

[163] Bertrand Jayles et al. How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences of the United States of America*, 114(47):12620 – 12625, 2017.

Appendices

Computation of Error Bars

The error bars should attempt to translate the variability of our results depending on the N_q questions presented to the subjects. In order to do so, we have defined a kind of bootstrap procedure for each point plotted.

We call x_0 the actual measurement of a quantity appearing in these figures by considering all N_q questions. Then, we generate the results of $N = 10000$ new effective experiments. For each effective experiment indexed by $n = 1, ..., N$, we randomly draw $Q = N_q$ questions among the N_q questions actually asked (so that some questions can appear several times, and others may not appear) and recompute the quantity of interest which now takes the value x_n. The upper error bar b_+ for x_0 is defined so that $C = 70\,\%$ of the x_n greater than x_0 are between x_0 and $x_0 + b_+$. Similarly, the lower error bar b_- is defined so that $C = 70\,\%$ of the x_n lower than x_0 are between $x_0 - b_-$ and x_0. The introduction of these upper and lower confidence intervals is adapted to the case were the distribution of the x_n is not symmetric, which is expected when the measured quantity x_0 is intrinsically positive and potentially close to 0.

Note that when x_0 is the *average* of a quantity (like in Figure 1.11), and if the law of large number applies (which should be marginally the case for N_q large enough), the x_n would be symmetrically and Gaussian distributed around x_0. Our error bars would then be symmetric ($b_+ = b_-$) and would be only slightly larger than a standard error, which corresponds to $C \approx 68.3\,\%$ for a Gaussian distributed random variable, instead of $C = 70\,\%$. We have checked that for figures where this is the case, our procedure indeed almost leads to identical error bars as the standard error $b = \lim_{N\to+\infty} \sum_{n=1}^{N}(x_0 - x_n)^2/N = \sigma/\sqrt{Q}$. σ is the usual standard deviation from question to question: $\sigma^2 = \sum_{q=1}^{Q}(x_0 - X_q)^2/Q$, where X_q is the quantity of interest measured for each question q.

Our procedure is particularly well suited for computing error bars for Figures 1.13 and 1.15, for which x_0 is a median or an average absolute deviation from the median, considering that there exists no robust definition of a standard error for a such quantities when the x_n are not Gaussian distributed. In particular, the fact that the measured quantities are intrinsically positive generally produces smaller lower error bars than upper error bars ($b_+ > b_-$) in these figures.

Sensitivity to Social Influence, Accuracy and Confidence

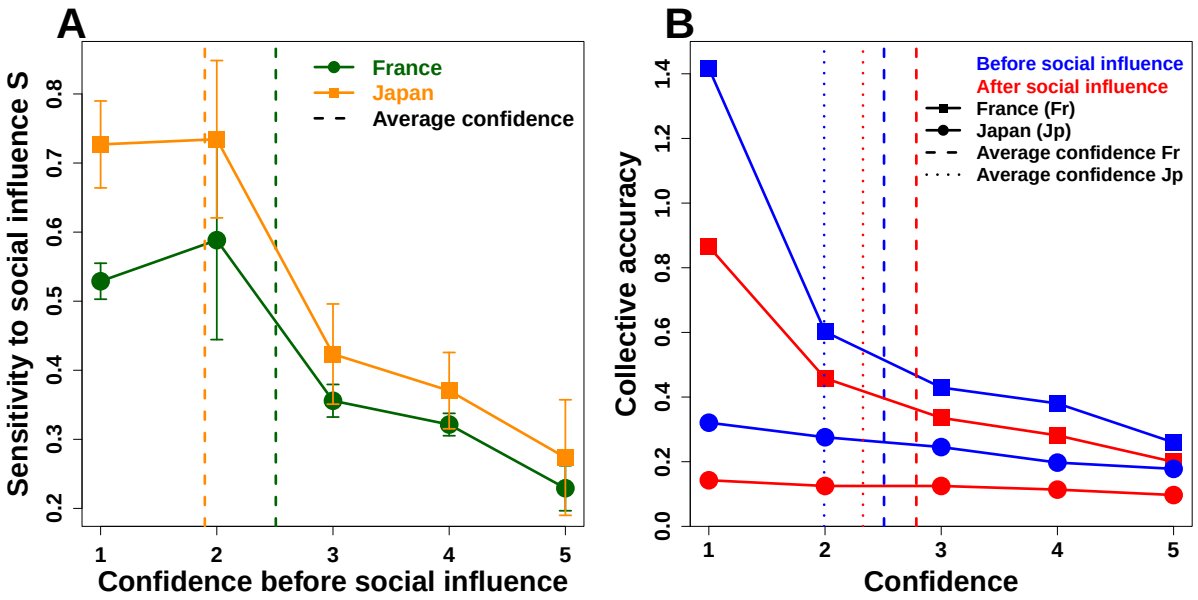

Figure 4.1: A. Average sensitivity to social influence S against the confidence reported in France (green) and Japan (orange). S is negatively correlated to confidence, meaning that the most confident subjects are also the less sensitive to social influence. The average sensitivity to social influence is higher in Japan than in France, and conversely, the average confidence is higher in France than in Japan. This is all the more surprising, because the questions were "harder" (lower demonstrability) in France. This suggests cultural differences in expression of confidence and attention to others' opinions, as already underlined in Chapter 1 sections 1.4.1 and 1.4.3. B. Collective accuracy (median of absolute values of log-transformed estimates), before (blue) and after social influence (red), for the experiments performed in France (squares) and Japan (circles). Collective accuracy is positively correlated with confidence, meaning that the most confident individuals are also the most accurate in their answers (answers closest to 0), arguably because they had a better prior knowledge. As explained in the main text (subsection 1.6.2), a complementary explanation is that individuals whose personal estimate falls close to the social information – and who hence tend to keep their opinion more – are more accurate on average because the social information itself is on average more accurate than random estimates. The average accuracy is better in Japan than in France, because of the difference in the questions' "difficulty" (higher demonstrability in the experiment in Japan). The average confidence increases and the accuracy improves after social influence.

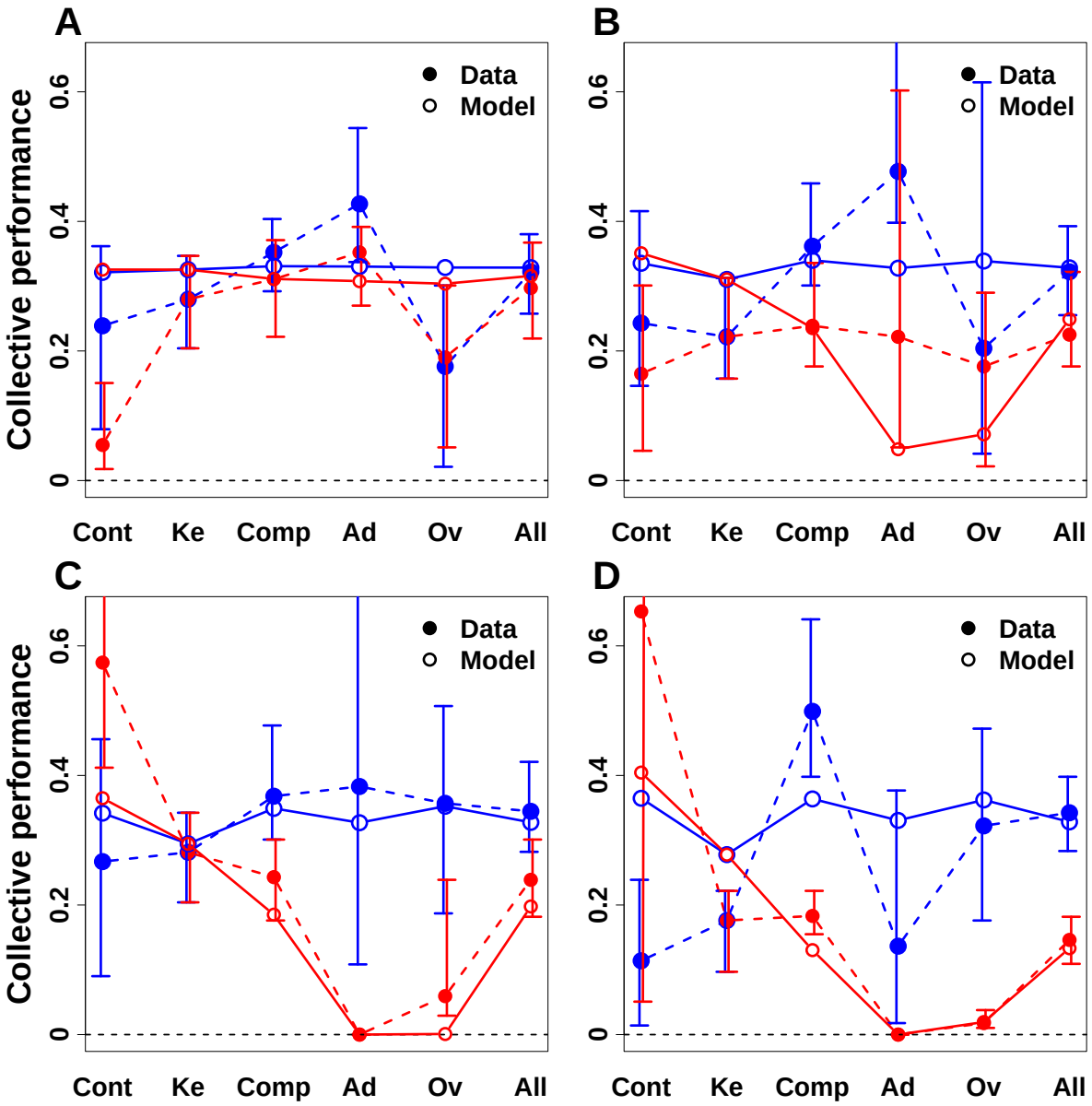

Figure 4.2: Collective performance (absolute value of the median of log-transformed estimates), for the different values of ρ, before (blue) and after (red) social influence, in the first experiment performed in France: A. $\rho = 0\%$; B. $\rho = 20\%$; C. $\rho = 43\%$; D. $\rho = 80\%$. The results for the five behavioral categories identified in Figure 1.8B, as well as the overall group (All), are presented. Experimental values correspond to full circles, and simulation values to empty circles. Similarly to the collective accuracy, the behavior leading to the best improvement in performance consists in adopting others' opinion, once virtual experts are introduced. Again, the improvement in collective performance after social influence increases with ρ.

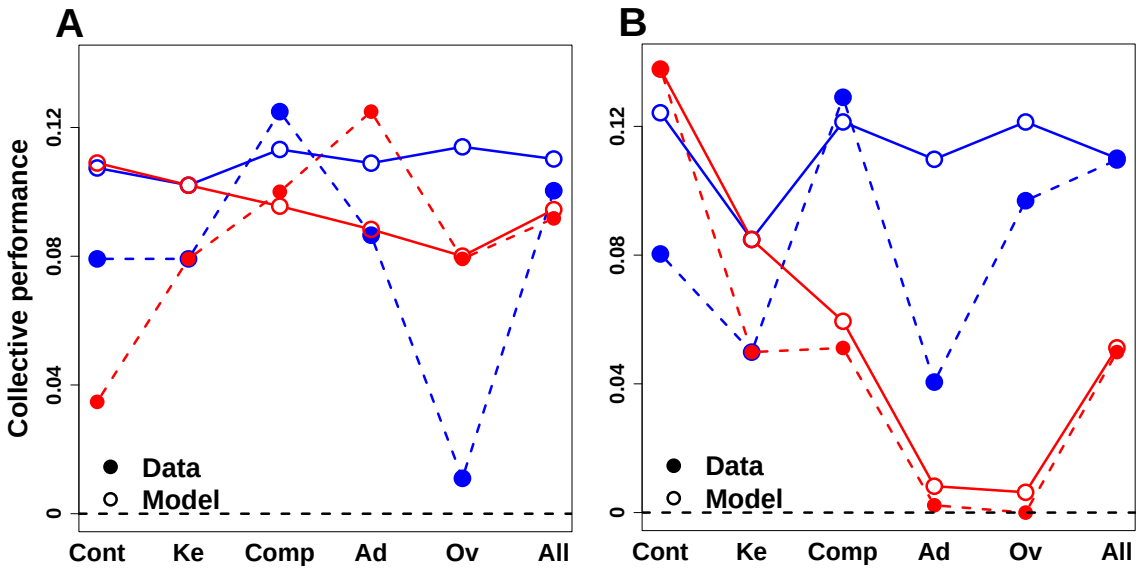

Figure 4.3: Collective performance (absolute value of the median of log-transformed estimates) for the first experiment performed in Japan, before (blue) and after social influence (red), for the two values of ρ (percentage of experts): 0% (A) and 33% (B). Adopting and overreacting also lead to the best improvement in collective performance after social influence. The improvement seems to increase with ρ, although we cannot assert it, because the questions asked in A and B were different.

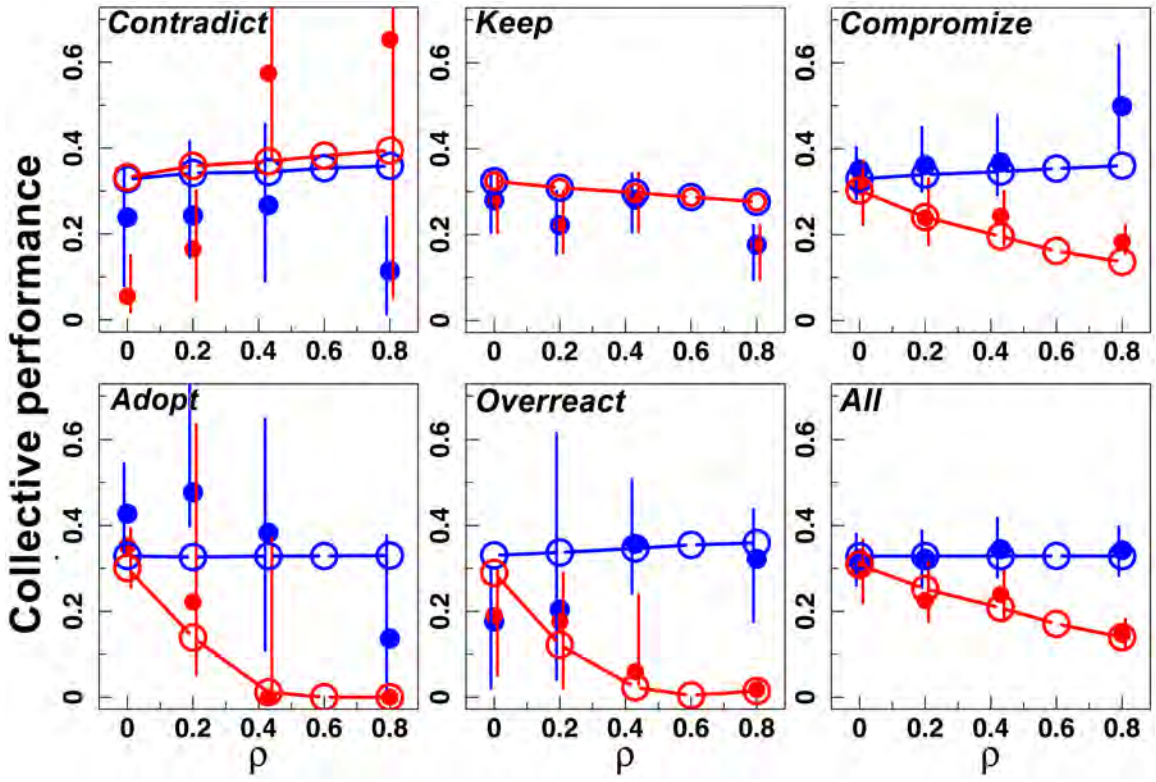

Figure 4.4: Collective performance (absolute value of the median of log-transformed estimates) before (blue) and after (red) social influence against ρ, for the 5 behavioral categories identified in Figure 1.8B and for the whole group (All). Adopting leads to the sharpest improvement, and the best performance for $\rho \geq 40\,\%$. Full circles correspond to experimental data, while empty circles represent the predictions of the model (including for $\rho = 60\,\%$, a case not tested experimentally).

Extraction Procedure

The following procedure allows to find the experimental form of the functions used in the equations of motion. So far it was run for one pedestrian only, but a lot of information can already be extracted.

In polar coordinates, the velocity equation – for one pedestrian – reads[5]:

$$\frac{d\vec{v}}{dt} = -A(v)\vec{e}_{\parallel} - B(v)f(r_{\mathrm{w}})g(\theta_{\mathrm{w}})\cos(\theta_{\mathrm{w}})\vec{e}_{\parallel} + B(v)f(r_{\mathrm{w}})g(\theta_{\mathrm{w}})\sin(\theta_{\mathrm{w}})\vec{e}_{\perp} + \sigma_{\parallel}\vec{\eta}_{\parallel} + \sigma_{\perp}\vec{\eta}_{\perp}. \quad (4.22)$$

The derivative of $\vec{v} = v\,\vec{e}_{\parallel}$ can also be written:

$$\frac{d\vec{v}}{dt} = \frac{dv}{dt}\,\vec{e}_{\parallel} + v\,\frac{d\vec{e}_{\parallel}}{dt} = \frac{dv}{dt}\,\vec{e}_{\parallel} + v\,\frac{d\phi}{dt}\,\vec{e}_{\perp}, \quad (4.23)$$

where ϕ is the angle between $\vec{v}$ and the horizontal (see Figure 2.7)
Identifying equations 4.22 and 4.23, we have:

$$\frac{dv}{dt} = -A(v) - B(v)f(r_{\mathrm{w}})g(\theta_{\mathrm{w}})\cos\theta_{\mathrm{w}} + \sigma_{\parallel}\eta_{\parallel}, \quad (4.24)$$

$$v\frac{d\phi}{dt} = B(v)f(r_{\mathrm{w}})g(\theta_{\mathrm{w}})\sin\theta_{\mathrm{w}} + \sigma_{\perp}\eta_{\perp}. \quad (4.25)$$

Discretization. We discretize the three variables r_{w}, θ_{w} and v in N_r, N_θ and N_v intervals respectively:

- $\{r_{\mathrm{w_i}}\}_{i=1}^{N_r+1}$, $r_{\mathrm{w_i}} = \frac{i-1}{N_r}\,R$,
- $\{\theta_{\mathrm{w_j}}\}_{j=1}^{N_\theta+1}$, $\theta_{\mathrm{w_j}} = \frac{j-1}{N_\theta}\,2\pi$,
- $\{v_k\}_{k=1}^{N_v+1}$, $v_k = \frac{k-1}{N_v}\,(V_{\max} - V_{\min})$,

where R is the radius of the arena, $V_{\max}$ and $V_{\min}$ are chosen such that there is enough data in each box. We thus obtain a discrete 3-dimensional space for the 3 variables.

Then, at each time t, we need to evaluate $\frac{dv(t)}{dt}$ and $v\frac{d\phi(t)}{dt}$, as well as $r_{\mathrm{w}}(t)$, $\theta_{\mathrm{w}}(t)$ and $v(t)$), which define the pedestrian's state at this time. The triplet $(r_{\mathrm{w}}(t), \theta_{\mathrm{w}}(t), v(t))$ maps to a 3D box ijk in the 3D space above defined. In each box, we average dv/dt and $vd\phi/dt$, and note the results q_{ijk} and p_{ijk} respectively[6].

We further assume that in each box, the average values of $A(v)$, $B(v)$, $f(r_{\mathrm{w}})$ and $g(\theta_{\mathrm{w}})$ (to be determined) can be approximated by their values in the center of the box, that we note A_k, B_k, f_i and g_j. We thus have:

$$q_{ijk} = -A_k - B_k\,f_i\,g_j\,\cos\theta_{\mathrm{w_j}}, \quad (4.26)$$

$$p_{ijk} = B_k\,f_i\,g_j\,\sin\theta_{\mathrm{w_j}}. \quad (4.27)$$

[5] We hide time dependence notations for clarity.

[6] Our data cover almost 120 minutes of random walk for each arena size, with a time step of 0.1 s (after smoothing the raw data, originally with a time step of 0.5 s). This large amount of data yields accurate averages.

Error minimization. To determine the values of $\{f_i\}_{i=1}^{N_r}$, $\{g_j\}_{j=1}^{N_\theta}$, $\{A_k\}_{k=1}^{N_v}$ and $\{B_k\}_{k=1}^{N_v}$ in Eqs. 4.26 and 4.27, we define an error function Δ to minimize:

$$\Delta = \sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta}\sum_{k=1}^{N_v} \epsilon_{ijk}\left[\left(q_{ijk} + A_k + B_k\, f_i\, g_j \cos\theta_{\mathrm{w_j}}\right)^2 + \left(p_{ijk} - B_k\, f_i\, g_j \sin\theta_{\mathrm{w_j}}\right)^2\right]. \tag{4.28}$$

Minimizing Δ amounts to search the solutions of $\dfrac{\partial\Delta}{\partial A_k} = \dfrac{\partial\Delta}{\partial B_k} = \dfrac{\partial\Delta}{\partial f_i} = \dfrac{\partial\Delta}{\partial g_j} = 0$.

- Equation for A_k:

$$\frac{\partial\Delta}{\partial A_k} = \sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk}\frac{\partial}{\partial A_k}\left(q_{ijk} + A_k + B_k\, f_i\, g_j \cos\theta_{\mathrm{w_j}}\right)^2 \tag{4.29}$$

$$= 2\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk}\left(q_{ijk} + A_k + B_k\, f_i\, g_j \cos\theta_{\mathrm{w_j}}\right) = 0, \tag{4.30}$$

$$A_k = -\frac{\displaystyle\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta}\epsilon_{ijk}\left(q_{ijk} + B_k\, f_i\, g_j \cos\theta_{\mathrm{w_j}}\right)}{\displaystyle\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta}\epsilon_{ijk}}, \quad \text{for } k = 1, \ldots, N_v. \tag{4.31}$$

- Equation for B_k:

$$\begin{aligned}
\frac{\partial\Delta}{\partial B_k} &= \sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk}\frac{\partial}{\partial B_k}\left[\left(q_{ijk} + A_k + B_k f_i g_j \cos\theta_{\mathrm{w_j}}\right)^2 + \left(p_{ijk} - B_k f_i g_j \sin\theta_{\mathrm{w_j}}\right)^2\right] \\
&= 2\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk}\left[\left(q_{ijk} + A_k + B_k f_i g_j \cos\theta_{\mathrm{w_j}}\right)\left(f_i g_j \cos\theta_{\mathrm{w_j}}\right) + \left(p_{ijk} - B_k f_i g_j \sin\theta_{\mathrm{w_j}}\right)\left(-f_i g_j \sin\theta_{\mathrm{w_j}}\right)\right] \\
&= 2\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk}\left[(q_{ijk} + A_k) f_i g_j \cos\theta_{\mathrm{w_j}} + B_k (f_i g_j \cos\theta_{\mathrm{w_j}})^2 - p_{ijk} f_i g_j \sin\theta_{\mathrm{w_j}} + B_k (f_i g_j \sin\theta_{\mathrm{w_j}})^2\right] \\
&= 2\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk}\left[B_k (f_i g_j)^2 + (q_{ijk} + A_k) f_i g_j \cos\theta_{\mathrm{w_j}} - p_{ijk} f_i g_j \sin\theta_{\mathrm{w_j}}\right] = 0, \\
\text{so} \qquad & \sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk} B_k (f_i g_j)^2 = \sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta} \epsilon_{ijk} f_i g_j \left[p_{ijk}\sin\theta_{\mathrm{w_j}} - (q_{ijk} + A_k)\cos\theta_{\mathrm{w_j}}\right], \\
\textit{i.e.} \qquad & B_k = \frac{\displaystyle\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta}\epsilon_{ijk} f_i g_j \left[p_{ijk}\sin\theta_{\mathrm{w_j}} - (q_{ijk} + A_k)\cos\theta_{\mathrm{w_j}}\right]}{\displaystyle\sum_{i=1}^{N_r}\sum_{j=1}^{N_\theta}\epsilon_{ijk}(f_i g_j)^2}, \quad \text{for } k = 1, \ldots, N_v.
\end{aligned} \tag{4.32}$$

- Equations for f_i and g_j lead to the same expression as B_k, with B_k interchanged with f_i

(resp. g_j) and index i (resp. j) with index k in the sums:

$$f_i = \frac{\sum_{j=1}^{N_\theta}\sum_{k=1}^{N_v}\epsilon_{ijk}B_k g_j\left[p_{ijk}\sin\theta_{\mathrm{w_j}} - (q_{ijk}+A_k)\cos\theta_{\mathrm{w_j}}\right]}{\sum_{j=1}^{N_\theta}\sum_{k=1}^{N_v}\epsilon_{ijk}(B_k g_j)^2}, \quad \text{for } i = 1,\ldots,N_r, \tag{4.33}$$

$$g_j = \frac{\sum_{i=1}^{N_r}\sum_{k=1}^{N_v}\epsilon_{ijk}B_k f_i\left[p_{ijk}\sin\theta_{\mathrm{w_j}} - (q_{ijk}+A_k)\cos\theta_{\mathrm{w_j}}\right]}{\sum_{i=1}^{N_r}\sum_{k=1}^{N_v}\epsilon_{ijk}(B_k f_i)^2}, \quad \text{for } j = 1,\ldots,N_\theta. \tag{4.34}$$

Successive Over-Relaxation (SOR) process. SOR is a method to solve complex systems of equations with a converging iterative process, by writing the variables x (x is here A_k, B_k, f_i or g_j) as follows:

Equations 4.31 to 4.34 can be written as $\mathbf{x} = F(\mathbf{x})$, where $\mathbf{x} = (A_k, B_k, f_i, g_j)$. It is a fixed point equation, which can be solved using a fixed-point iteration: $\mathbf{x}^{l+1} = f(\mathbf{x}^l)$, where l is the iteration step. We use the SOR method, writing the fixed point equation as:

$$\mathbf{x}^{(l+1)} = (1-\lambda)\mathbf{x}^{(l)} + \lambda\mathbf{x}^*, \tag{4.35}$$

where l is the iteration index, λ is the relaxation parameter, and $\mathbf{x}^*$ is the value of $\mathbf{x}$ calculated from equations 4.31 to 4.34 with the values of the variables at step l. If this process converges, then the convergence value is the solution of the fixed-point equation. Starting from reasonable initial values of the parameters – we used $A_k^{(0)} = B_k^{(0)} = 1\ \forall k$, $f_i^{(0)} = 1\ \forall i$ and $g_j^{(0)} = 1\ \forall j$ – we found that the process converges in less than 50 iteration for $\lambda = 0.25$. There could be several fixed-points for a given set of initial conditions, but we checked that this was not the case.

Normalization at each iteration. In Eqs. 4.26 and 4.27, the functions B, f and g appear always as a factor Bfg, such that they are defined to a constant. We are free to choose normalization conditions, and propose that $B(v)$ must satisfy:

$$\int_{V_{\min}}^{V_{\max}} B(v)\rho(v)\,dv = \int_{V_{\min}}^{V_{\max}} \rho(v)\,dv,$$

i.e., $\{B_k\}_{k=1}^{N_v}$ must be such that (after removing the factor $(V_{\max}-V_{\min})/K$ at both sides)

$$\sum_{k=1}^{N_v} B_k\rho_k = \sum_{k=1}^{N_v}\rho_k,$$

and that g must satisfy:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} g^2(\theta_{\mathrm{w}})d\theta_{\mathrm{w}} = 1. \tag{4.36}$$

Let us introduce the following notation:

$$S_B = \sum_{k=1}^{N_v} B_k^{(l+1)}\rho_k, \quad S_\rho = \sum_{k=1}^{N_v}\rho_k, \quad S_g = \sqrt{\frac{1}{J}\sum_{j=1}^{N_\theta}\left(g_j^{(l+1)}\right)^2}. \tag{4.37}$$

Then, the normalization is

$$B_k^{\mathrm{new}} = B_k^{(l+1)}S_\rho/S_B, \quad g_j^{\mathrm{new}} = g_j^{(l+1)}/S_g \quad \text{and} \quad f_i^{\mathrm{new}} = f_i^{(l+1)}S_B S_g/S_\rho. \tag{4.38}$$

Résumé Détaillé en Français

Introduction

L'information tient une place prépondérante dans l'aventure humaine, en ce qu'elle relie les développements concomitants de la technologie et des moyens de communication. Dans les sociétés – notamment occidentales – contemporaines, et en particulier dans les dernières décennies qui ont vu émerger Internet et les réseaux sociaux (Facebook, Twitter...), la quantité d'information à laquelle sont soumis les individus dépasse de loin leurs capacités individuelles de traitement et d'intégration [5].

De plus, la défiance grandissante vis-à-vis des médias traditionnels, couplée aux capacités de diffusion massive de l'information des réseaux sociaux, offre aux informations frauduleuses ou non-vérifiées la possibilité de se propager avec une ampleur sans précédent [36, 37, 44]. Cette situation amène de nombreuses questions: doit-on contrôler ce déluge d'informations auquel chacun est soumis? Si oui, comment? Comment définir l'information pertinente pour un individu particulier? Quelle quantité d'information un individu peut-il traiter? Comment distinguer une information viable d'une fake news?...

La complexité immense de ces problèmes amène à se pencher vers des solutions collectives: de nombreux espoirs reposent sur les recherches en *intelligence collective*, cette capacité des groupes, observée dans certaines circonstances, à résoudre des problèmes de façon plus efficace qu'un individu isolé [6, 7, 8]. L'intelligence collective est un domaine de recherche extrêmement vaste, auquel on peut s'attaquer de différentes manières. Notre recherche s'appuie sur deux des approches les plus prometteuses en la matière: *la sagesse des foules* [9] et *l'intelligence d'essaim* [10].

Dans cette thèse, nous nous intéressons aux conditions sous lesquelles des interactions contrôlées entre les individus d'un groupe humain peuvent conduire celui-ci à trouver ou à se rapprocher de la bonne solution à un problème. En particulier, nous attaquons de front le problème de l'impact de la *quantité* et de la *qualité* de l'information à laquelle les individus d'un groupe ont accès, sur leur capacité à résoudre collectivement certains types de tâches. Nous abordons aussi la question du développement d'outils de traitement de l'information, dans une optique d'aide à la prise de décision.

Pour cela, nous avons mené deux types d'expériences, où les sujets devaient accomplir des tâches bien distinctes, et qui constituent les deux projets majeurs développés dans cette thèse:

1. Dans le premier type d'expériences, les sujets devaient estimer des quantités diverses (âges, nombre de billes dans des jarres, populations de villes, questions astronomiques...), puis, après avoir reçu comme information sociale la moyenne des estimations d'autres individus sur la même question, avaient la possibilité de réviser leur jugement.

2. Dans le second type d'expériences, des groupes de 22 piétons devaient marcher à l'intérieur

d'un cercle, puis se séparer en clusters de la même couleur, avec pour seule information un signal sonore provenant de tags attachés à leur épaules, ce signal étant émis dans des conditions bien spécifiques.

Nous allons maintenant exposer les principaux résultats obtenus dans chacun de ces projets.

Impact de la Quantité et de la Qualité de l'Information sur la Précision des Estimations

Nous avons mené en tout quatre expériences, deux au Japon et deux en France. En plus de vérifier la validité de nos résultats dans deux cultures bien différentes sur de nombreux aspects, ces expériences sont aussi différentes étapes d'un long programme de recherche combinant étroitement modélisation théorique, analyse de données et expériences.

Introduction

Les tâches d'estimation constituent un cadre très pratique pour faire des mesures quantitatives. Elles sont le cadre de référence des travaux sur la *sagesse des foules*, théorie basée sur l'hypothèse que la connaissance dans un groupe est distribuée, et qu'une aggrégation adéquate de cette connaissance peut mener un groupe à prendre collectivement de meilleures décisions que les individus isolés, experts compris.

Le premier et plus célèbre exemple de ce phénomène est l'expérience de Galton en 1907 [47]. Galton a proposé aux participants d'une foire d'estimer le poids d'un boeuf, et trouva que l'estimation médiane tombait à moins d'un pourcent de la vraie valeur, ce qui était mieux que les meilleures estimations individuelles. Galton appela ce phénomène *Vox populi*, et James Surowiecki popularisa l'expression *Sagesse des foules* en 2005 dans son best-seller éponyme [9].

Des travaux récents ont suggéré que les bénéfices de ce phenomène se trouvaient atténués si les individus du groupe étaient autorisés à interagir, car cela diminuait la diversité des estimations [58]. Cependant, ces résultats proviennent de définitions de la performance d'un groupe dont nous mettons en question la pertinence. D'ailleurs, des travaux ultérieurs ont montré que l'influence sociale peut améliorer les performances d'un groupe si correctement utilisée, par exemple en repérant les individus informés ou confiants [63, 60].

Des considérations purement mathématiques mettent aussi en doute ce résultat: en effet, si l'on note E_i l'estimation d'un individu i, et T la vraie valeur à trouver, alors $\mathscr{G}_D = \langle (E_i - \langle E_i \rangle)^2 \rangle$ est une mesure naturelle de la diversité du groupe, $\mathscr{G} = (\langle E_i \rangle - T)^2$ est une mesure de la précision collective du groupe, et $\mathscr{G}' = \langle (E_i - T)^2 \rangle$ est une mesure de la précision individuelle moyenne. Ces trois mesures (communément utilisées dans la litérature) sont liées par la relation suivante: $\mathscr{G}' = \mathscr{G} + \mathscr{G}_D$, qui montre qu'une diminution de la diversité $\mathscr{G}_D$ améliore la précision individuelle moyenne $\mathscr{G}'$.

Dans ce qui suit, nous montrerons que l'information sociale aide effectivement à améliorer la precision individuelle et collective dans des tâches d'estimation, et ce d'autant plus que l'information disponible est de qualité.

Dispositif Expérimental

La plupart des travaux utilisant ce type de tâche ont utilisé une transmission globale de l'information, fournissant l'information sociale à l'ensemble des individus d'un groupe en même temps. Mais il a été montré que dans un groupe, les décisions sont généralement prises séquentiellement

plutôt que simultanément [113, 114], ce qui nous a conduit à définir des tâches d'estimations séquentielles pour nos expériences.

Comme expliqué dans l'introduction, un individu i, après avoir proposé une première estimation E_{p_i} d'une quantité donnée (nous appellerons cela son estimation personnelle), reçoit de l'information sociale (c'est-à-dire de la part d'autres participants) I_i, et peut ensuite proposer une nouvelle estimation E_i. Puisque nous les humains pensons en ordres de grandeur [121], nous avons choisi de définir l'information sociale comme la moyenne géométrique des τ précédentes estimations ($\tau = 1, 3$): $I_i = (\prod_{j=i-\tau}^{i-1} E_j)^{\frac{1}{\tau}}$. L'estimation finale de l'individu E_i (après avoir reçu l'information sociale donc) est un compromis entre son estimation personnelle et l'information sociale: $E_i = {E_{p_i}}^{1-S_i} {I_i}^{S_i}$. Nous notons S_i le poids accordé à l'information sociale, et l'appelons *sensibilité à l'influence sociale*. $S_i = 0$ signifie que l'individu i garde son opinion, tandis que $S_i = 1$ signifie que l'individu i s'aligne sur l'information sociale.

D'autre part, en introduisant dans la séquence d'estimations des agents virtuels fournissant une valeur T_I de notre choix, nous avions la possibilité de contrôler précisément la quantité (proportion ρ d'agents) et la qualité (valeur de T_I) de l'information échangée par les sujets.

Distribution des Estimations et Information Préalable

Ainsi que mentionné dans l'introduction, l'un des enjeux majeurs de la recherche sur la *sagesse des foules* est de trouver le bon moyen d'aggréger les estimations individuelles de façon à en "extraire" la "sagesse" collective. La question devient encore plus épineuse lorsque l'on souhaite aggréger des estimations de quantités différentes.

En effet, lorsque les quantités à estimer diffèrent largement, notamment de plusieurs ordres de grandeur, il convient de normaliser les estimations par la vraie valeur correspondante pour pouvoir les comparer [59]:

$$E_{\mathrm{norm}_{i,q}} = \frac{E_{i,q}}{T_q} \tag{5.39}$$

Cela ne vaut évidemment que dans les cas où l'on connaît la vraie valeur et où l'on s'intéresse par exemple à l'influence de l'information sociale (ce qui est notre cas).

De plus, il a été montré que, dès que les quantités à estimer sont suffisamment grandes (2 à 3 ordres de grandeur minimum), leur distribution est largement biaisée vers la droite, et que l'on peut symmétriser ces distributions en considérant le logarithme des estimations [58, 117]:

$$X_{i,q} = \log(\frac{E_{i,q}}{T_q}) \tag{5.40}$$

Notons que ceci est en adéquation avec le fait que les humains pensent essentiellement les nombres en termes d'ordres de grandeur. Les quantités naturelles à considérer dans les tâches d'estimations sont donc les logarithmes des estimations, plutôt que les estimations elles-mêmes.

De plus, les distributions usuelles étant caractérisées par leur centre m et leur largeur σ, et ces paramètres variant fortement selon la quantité à estimer, il convient de les prendre en compte dans le processus de normalisation:

$$Z_{i,q} = \frac{X_{i,q} - m_q}{\sigma_q} \tag{5.41}$$

La Figure 5.1 montre les distributions de cette quantité pour les 4 expériences menées.

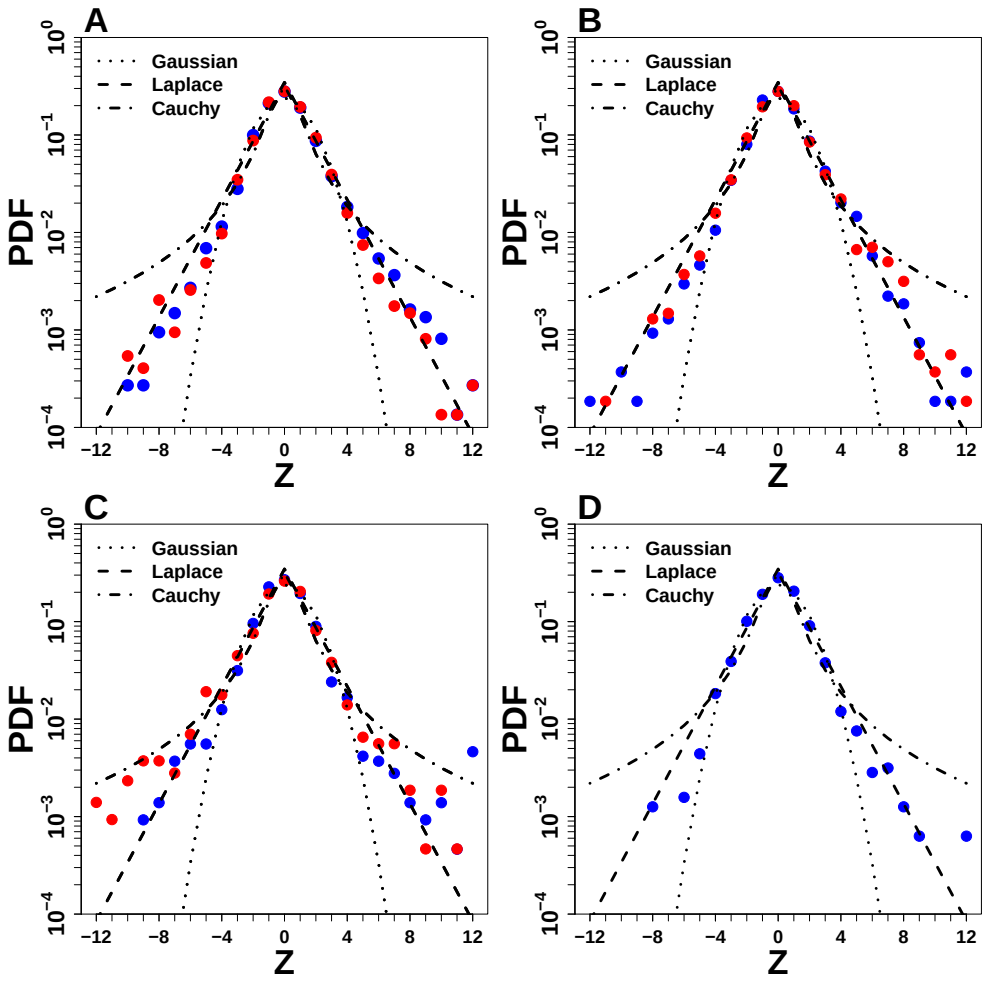

Figure 5.1: Distribution des estimations individuelles normalisées en A. France en 2017; B. France en 2016; C. Japon en 2015; Japon en 2016, avant (bleu) et après (rouge) influence sociale. Les lignes noires sont les distributions standard (centre 0 et largeur 1) de Cauchy, Gauss et Laplace. Les points expérimentaux sont bien mieux fittés par une distribution de Laplace que par une distribution Gaussienne ou de Cauchy.

Nous obtenons de belles distributions de Laplace, et non Gaussiennes comme généralement proposé dans la litérature. Des recherches plus approfondies nous ont permis de comprendre l'origine de cette différence, et de proposer par la même occasion un lien entre la quantité d'information préalablement détenue par un groupe et la distribution attendue de leurs estimations.

Tout d'abord, les distributions Gaussienne et de Laplace appartiennent à la famille des Distributions Normales Généraliseées (DNG), de PDF [128]:

$$f(X, m, \sigma, n) = \frac{1}{2\,\sigma\,\Gamma(1+1/n)} \exp\left\{-\left|\frac{X-m}{\sigma}\right|^n\right\} \tag{5.42}$$

où n est le paramètre de "forme" (qui définit les "queues" des distributions): $n = 1$ pour une distribution de Laplace, et $n = 2$ pour une distribution Gaussienne. Une étude sur de nombreux jeux de données a montré que les distributions d'estimations sont effectivement des DNG, mais généralement plus proches de Laplaces que de Gaussiennes [129].

Reste à savoir l'origine de ces différences, et en particulier pourquoi dans nos expériences, les distributions sont tellement plus proches de Laplaces que de Gaussiennes. L'une des différences

majeures entre notre étude et les études préalables, est que nous avons choisi des quantités "difficiles", dans le sens où l'on s'attendait à ce que les participants n'aient qu'une idée très vague de la vraie valeur à estimer. Ceci nous a mis sur la piste de l'information préalable des participants quant à une quantité donnée à estimer: plus les participants ont d'information préalable sur une quantité à estimer (c'est-à-dire plus cette quantité est "facile"), plus la distribution de leurs estimations sera proche d'une Gaussienne; inversement, moins les participants ont d'information préalable sur une quantité donnée à estimer (c'est-à-dire plus cette quantité est "difficile"), plus la distribution de leurs estimations sera proche d'une Laplace.

L'argument est le suivant: la distribution de Laplace maximise la distribution de probabilité d'entropie sous la contrainte que son premier moment (déviation moyenne absolue de la médiane σ) est fixé [130]. Cela signifie que pour une valeur de σ donnée, et si rien d'autre n'est connu à propos de la distribution, le plus probable est que l'on observe une distribution de Laplace.

On peut faire un parallèle direct avec la distribution de Boltzmann, qui donne la distribution de probabilité de états d'énergie ϵ dans un système de particules à énergie fixe k_BT: $f(\epsilon, T) \propto e^{-\frac{\epsilon}{k_BT}}$, où k_B est la constante de Boltzmann et T est la température du système. La distribution exponentielle maximise la distribution de probabilité d'entropie pour une valeur fixe de l'énergie moyenne k_BT. Les états d'énergie ϵ jouent un rôle équivalent aux log-estimations X dans nos expériences, et l'énergie moyenne k_BT un rôle équivalent à σ.

De même, la distribution gaussienne maximise la distribution de probabilité d'entropie sous les contraintes que la moyenne (premier moment) et la variance (second moment) sont fixés. On comprend que cette contrainte supplémentaire sur le second moment équivaut à une quantité beaucoup plus importante d'information préalable dans le système (le groupe).

En d'autres termes, en posant une certaine question à un certain groupe, l'expérimentateur "fixe" certaines contraintes, qui dépendent de la quantité d'information préalable détenue par les individus du groupe sur la quantité à estimer. Si le groupe a très peu d'information préalable, on s'attend à une contrainte minimale, à savoir une contrainte sur le premier moment uniquement: on doit alors obtenir une distribution de Laplace. Si le groupe a un niveau de connaissance antérieure beaucoup plus élevé, on peut s'attendre à une contrainte supplémentaire sur le second moment (pas d'estimations extrêmement éloignées de la vraie valeur): on doit alors obtenir une distribution gaussienne.

Le fait que les estimations (plus exactement leurs logarithmes) suivent généralement des distributions de Laplace plutôt que des distributions gaussiennes a des implications en termes d'aggrégation des données. En effet, les estimateurs du centre et de la largeur des distributions Gaussienne et de Laplace étant différents, des définitions différentes des performances individuelles et collectives, ainsi que de la diversité, devront être utilisées selon que l'on observe une distribution gaussienne ou de Laplace. Ainsi, les estimateurs (maximum likelihood) du centre et de la largeur sont respectivement la moyenne et l'écart-type pour une distribution Gaussienne, et la médiane et la déviation absolue moyenne à la médiane pour une distribution de Laplace.

Les mesures communément utilisées pour la performance d'un groupe dans des tâches d'estimation sont les suivantes, et sont parfaitement adaptées pour des distributions Gaussiennes:

- l'erreur collective, que nous appellerons plutôt *performance collective*: $(\langle X_i \rangle - T)^2$
- l'erreur individuelle moyenne, que nous appellerons plutôt *précision collective*: $\langle (X_i - T)^2 \rangle$
- la diversité: $\langle (X_i - \langle X_i \rangle)^2 \rangle$

Pour des distributions de Laplace, il convient de faire les changements suivants:

- Performance collective: $|\text{median}(X_i)|$ (distance du centre de la distribution à la vraie valeur 0)
- Précision collective: $\text{median}(|X_i|)$ (distance médiane des estimations individuelles à la vraie valeur 0)
- Diversité: $\langle |X_i - \text{median}(|X_i|)| \rangle$, qui est aussi l'estimateur de la largeur

Pour conclure cette section, nous montrons sur la Figure 5.2 les distributions des log-estimations $X = \log(\frac{E}{T})$, qui serviront de base dans toute la suite du chapitre.

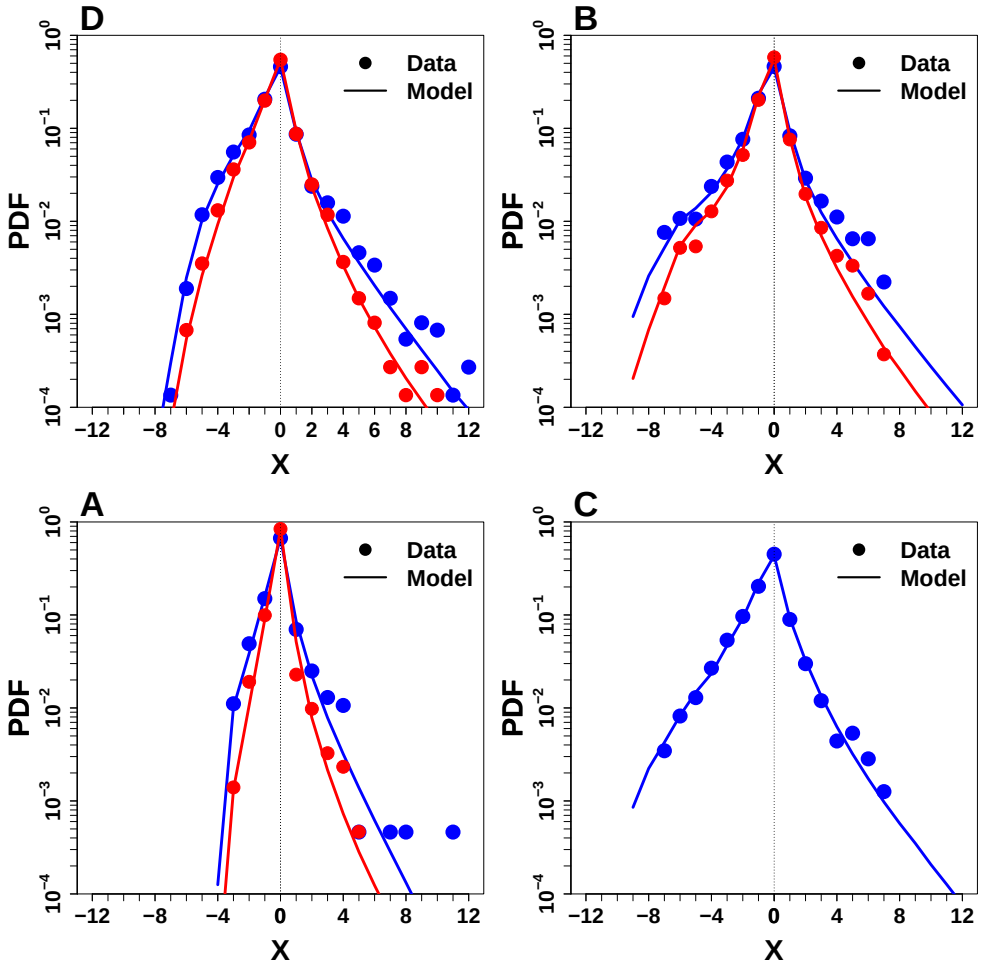

Figure 5.2: Distribution des (transformées logarithmiques des) estimations individuelles $X = \log(\frac{E}{T})$, avant (bleu) et après (rouge) influence sociale, pour l'expérience faite: A. en France en 2017; B. en France en 2016; C. au Japon en 2015; D. au Japon en 2016. Dans tous les cas, le modèle basé sur des distributions de Laplace (lignes continues) reproduit les données (points) de manière remarquable.

Notons que les distributions sont légèrement biaisées, soulignant la tendance humaine à sous-estimer les quantités, due à la représentation interne logarithmique des nombres [124]. Précisons également que la distribution des estimations au Japon en 2015 (panel C) est plus serrée que les autres. C'est parce que des questions faciles étaient mélangées aux questions difficiles, tandis que par la suite, seules des questions difficiles étaient posées. La seconde expérience au Japon visait simplement à vérifier que si l'on posait exactement les mêmes questions qu'en France, on obtiendrait alors une distribution très similaire, ce qui est le cas (voir Figure 5.3).

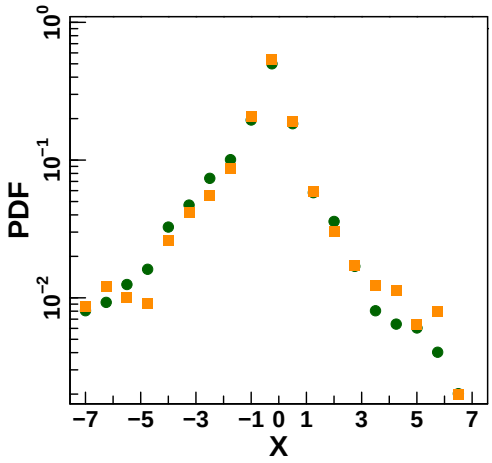

Figure 5.3: Fonction de distribution de probabilité (PDF) des estimations normalisées log-transformées $X = \log(\frac{E}{T})$, dans la première expérience en France (orange) et dans la deuxième expérience au Japon (vert). Les questions posées étaient les mêmes et les distributions sont très similaires, comme prévu.

Dans la section suivante, nous nous intéresserons à la sensibilité à l'influence sociale des individus.

Sensibilité à l'Influence Sociale

En termes de variables log-transformées $X_i = \log(\frac{E_i}{T})$, l'estimation après influence sociale s'écrit: $X_i = (1 - S_i)X_{p_i} + S_i M_i$, où l'information sociale log-transformée est simplement la moyenne arithmétique des τ estimations précédentes: $M_i = \frac{1}{\tau}\sum_{j=i-\tau}^{i-1} X_j$, ce qui nous permet de définir la sensibilité individuelle à l'influence sociale de manière unique:

$$S_i = \frac{X_i - X_{p_i}}{M_i - X_{p_i}} \tag{5.43}$$

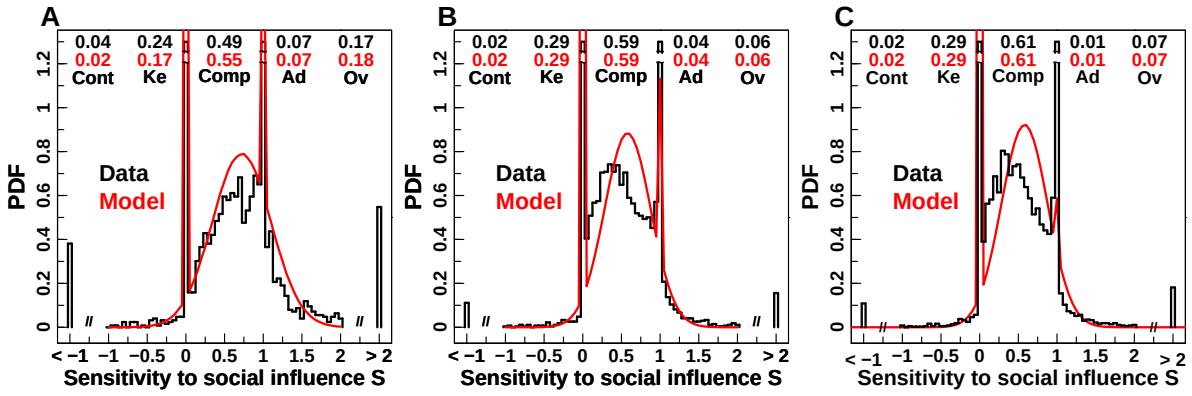

Figure 5.4: PDF des sensibilités à l'influence sociale S pour: A. la première expérience au Japon; B. la première expérience en France; C. la deuxième expérience en France (les sujets n'ont fourni que leurs estimations personnelles dans la seconde expérience au Japon, de sorte que la sensibilité à l'influence sociale n'a pas pu être mesurée). Les chiffres en haut de chaque figure sont les probabilités pour chaque catégorie de comportement: contredire l'information sociale (Cont; $S < 0$), garder son opinion (Ke; $S = 0$), faire un compromis (Comp; $0 < S < 1$), adopter l'information sociale (Ad; $S = 1$) et l'amplifier (Ov; $S > 1$). Les données expérimentales sont représentées en noir et les simulations numériques du modèle sont en rouge. La figure est limitée à l'intervalle $[-1, 2]$, et les valeurs de S en dehors de cet interval ont été regroupées dans les cases $S < -1$ et $S > 2$.

La Figure 5.4 montre les distributions de cette quantité pour les trois expériences menées (dans la seconde expérience au Japon, seules les estimations personnelles étaient récoltées).

Nous voyons une grande diversité dans les sensibilités à l'influence sociale S bien qu'un pattern général se dégage dans chaque distribution: deux pics à $S = 0$ et $S = 1$, et une tendance centrale pouvant grossièrement être assimilée à une gaussienne. Notons aussi que la distribution au Japon (Figure A) est décalée vers la droite par rapport à celles observées en France, ce qui indique une plus grande tendance à s'appuyer sur l'opinion des autres au Japon.

Ces distributions nous permettent de définir 5 *profils de comportement* bien distincts: conserver son opinion ($S = 0$), faire un compromis avec l'information sociale ($0 < S < 1$), adopter l'information sociale ($S = 1$), l'amplifier ($S > 1$) ou la contredire ($S < 0$). Selon leur degré de connaissance à une certaine question, les individus peuvent adopter l'un des 5 types de comportement.

Cependant, nous avons montré qu'il existe une certaine cohérence dans les comportements des individus, de sorte que nous avons identifié trois *profils de personnalité*: des individus "suiveurs", qui ont une grande tendance à utiliser l'information sociale tout au long des questions; des individus "confiants", qui ont au contraire une grande tendance à conserver leur opinion ou à en rester proche; et enfin des individus "moyens" qui ont plutôt tendance à faire des compromis. Les Figures 5.5 et 5.6 illustrent ce phénomène.

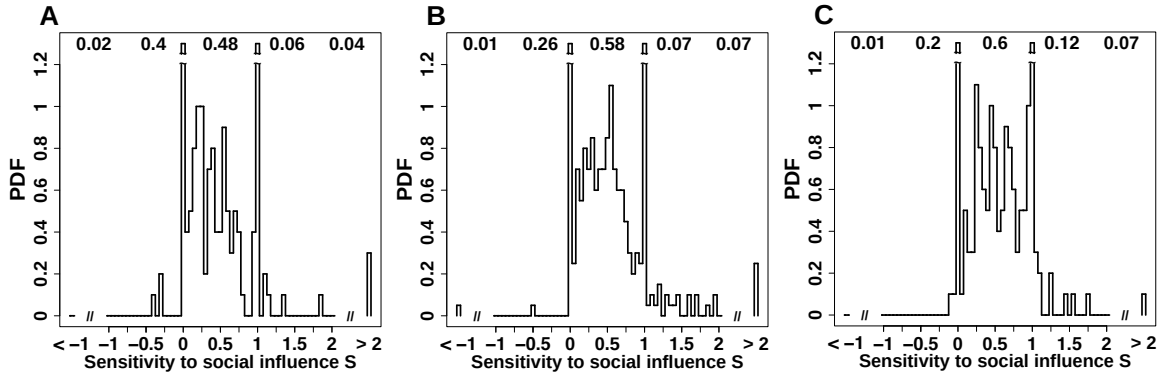

Figure 5.5: PDF de la sensibilité à l'influence sociale S, pour les questions 25 à 29 de la première expérience en France, dans les cas suivants: A. les sujets confiants, définis comme le quart de chaque sous-groupe de 8 sujets pour lesquels la quantité $\langle |S_q| \rangle_q$, où q est l'indice pour les questions 1 à 24, était minimisée (c'est-à-dire les sujets qui étaient en moyenne le plus proche de $S = 0$); B. les sujets "moyens", définis comme n'étant ni "confiants" ni "suiveurs"; C. les sujets suiveurs, définis comme le quart de chaque sous-groupe de 8 sujets pour lesquels la quantité $\langle |1 - S_q| \rangle_q$ était minimisée (c'est-à-dire les sujets qui étaient en moyenne le plus proche de $S = 1$). En A, le pic à $S = 0$ est presque 7 fois plus haut que celui à $S = 1$, tandis qu'en B, il est moins de 4 fois plus haut, et en C, moins de 2 fois. Ainsi, les sujets caractérisés comme confiants sur les questions 1 à 24 restent très confiants sur les questions 25 à 29, tandis que les sujets caractérisés comme suiveurs restent très suiveurs. De façon cohérente, les sujets identifiés comme "moyens" restent "moyens", dans le sens que la distribution de leurs sensibilités à l'influence sociale est très proche de la distribution globale (Figure 5.4B).

Nous avons ensuite identifié un phénomène non-trivial: la sensibilité à l'influence sociale des individus dépend en moyenne de la distance $D = X_p - M$ entre leur estimation personnelle (X_p) et l'information sociale (M). Plus l'information sociale s'éloigne de leur estimation personnelle, plus les sujets ont tendance à la suivre (Figure 5.7A). Il faut interpréter ce résultat comme une tendance accrue à douter de son opinion avec une plus grande divergence entre celle-ci et celle des autres. La Figure 5.7B montre d'ailleurs qu'à mesure que D augmente, la probabilité de garder son opinion diminue, au profit de celle de faire un compromis.

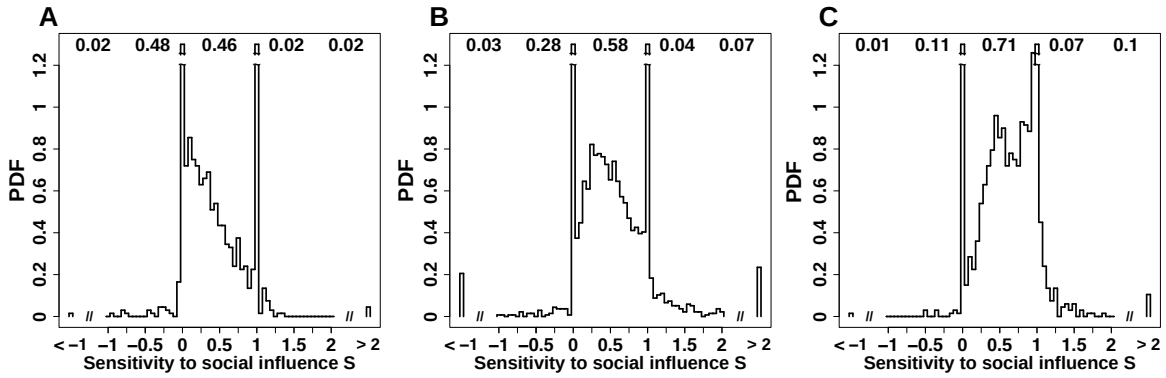

Figure 5.6: PDF de la sensibilité à l'influence sociale S, pour les individus identifiés comme confiants (A), "moyens" (B) ou suiveurs (C). Les trois catégories sont définies de la même façon qu'en Figure 5.5, mais cette fois le calcul a été fait sur les questions 1 à 29. Les différences apparaissent alors – naturellement – de façon plus claire qu'en Figure 5.5.

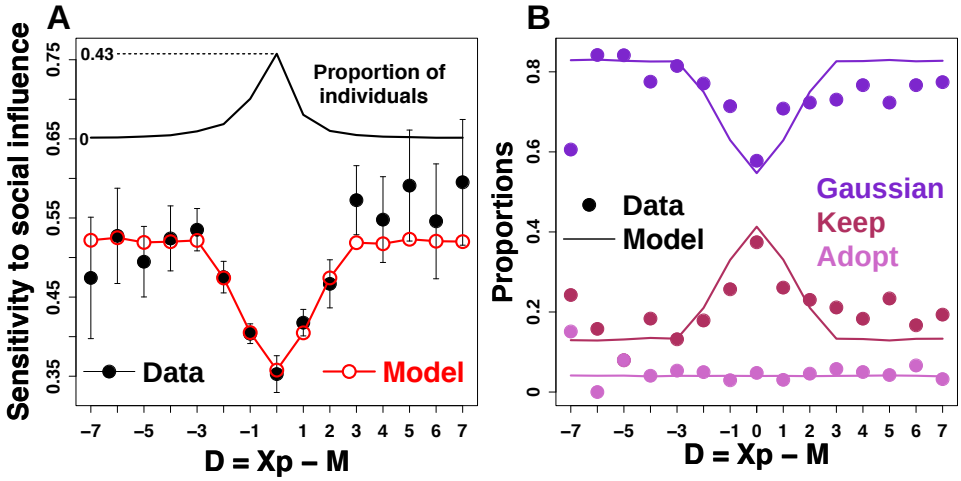

Figure 5.7: A. Sensibilité moyenne à l'influence sociale S en fonction de la distance entre estimation personnelle et information sociale $D = X_p - M$. Les points noirs correspondent aux données expérimentales, et les cercles rouges aux simulations du modèle. Notons qu'au-delà de 3 ordres de grandeur, il n'y a que 14 % des données; B. Fraction des sujets gardant leur opinion (bordeaux), adoptant l'information sociale (rose) et étant dans la partie gaussienne de la distribution des S (c'est-à-dire essentiellement des sujets faisant un compromis; violet) en fonction de D.

Tous ces éléments ont été incorporés dans notre modèle: les individus tirent leur estimation personnelle dans une distribution de Laplace, puis après avoir reçu l'information sociale, leur sensibilité à l'influence sociale S est définie d'après la relation linéaire observée en Figure 5.7A:

$$\langle S \rangle = P_0 \times 0 + P_1 \times 1 + P_g \times m_g = a + b\,|D|,$$

où a et la pente b sont extraits de la Figure 5.7A, et où P_0, P_1 et P_g sont respectivement les probabilités de garder son opinion, d'adopter celle des autres et d'être dans la partie Gaussienne de la distribution de S (voir Figure 5.4). P_1 étant un paramètre du modèle (la Figure 5.7B montre que P_1 est indépendant de D), on en déduit P_g en fonction de D puis $P_0 = 1 - P_1 - P_g$.

Nous avons ensuite comparé les prédictions du modèles avec les données expérimentales, puis utilisé le modèle pour étudier l'impact de la quantité et de la qualité de l'information sociale échangée sur la précision et la performance collectives, définies plus haut.

Impact de la Quantité d'Information sur la Précision des Estimations

Dans un premier temps, nous nous intéressons à de l'information correcte, et à l'impact de la quantité de celle-ci sur les performances du groupe dans des tâches d'estimation. Les agents virtuels, dont nous contrôlons la proportion ρ, fournissent alors la vraie valeur des quantités à estimer.

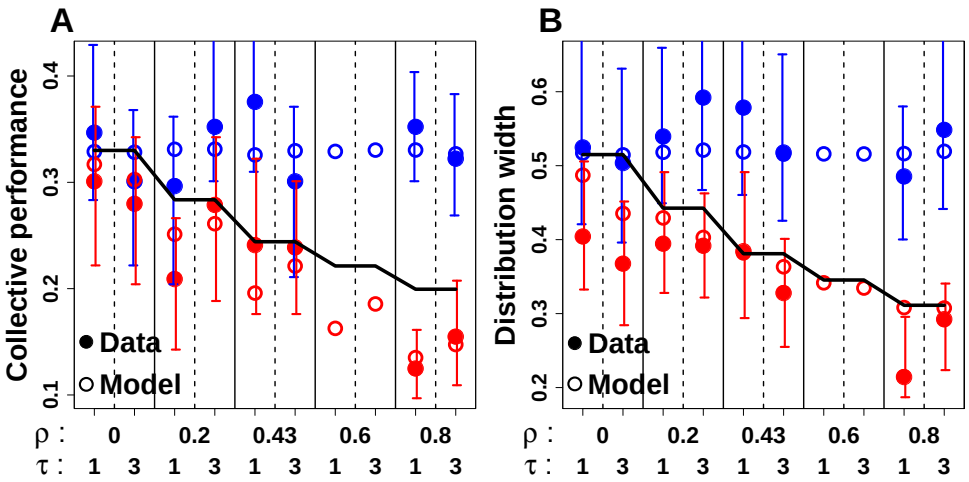

Figure 5.8: Performance collective (A; valeur absolue de la médiane des estimations) et largeur (B) de la distribution des (transformées logarithmiques des) estimations, pour tous les couples (ρ, τ), avant (bleu) et après (rouge) influence sociale. La performance collective augmente (c'est-à-dire le centre de la distribution se rapproche de la vraie valeur) et la distribution se resserre avec ρ, ainsi qu'après influence sociale. Les cercles pleins correspondent aux données expérimentales, tandis que les cercles vides correspondent aux prédictions du modèle. Les lignes continues noires sont les prédictions d'un modèle préliminaire, dans lequel S était indépendant de D. Pour $\rho = 60\,\%$, seules les prédictions du modèle sont disponibles.

La Figure 5.8 montre que le centre (médiane) de la distribution se rapproche de la vraie valeur (0 en coordonnées logarithmiques), et que la distribution se resserre, à mesure que la quantité d'information fournie aux sujets (à leur insu) augmente. Le rapprochement de la médiane de 0 signifie une amélioration de la performance collective, et comme la distribution se resserre, on s'attend aussi à ce que la précision collective s'améliore (par analogie avec la relation $\mathscr{G}' = \mathscr{G} + \mathscr{G}_D$ sur les moyennes explicitée plus haut).

La Figure 5.9 montre la précision collective du groupe avant (en bleu) et après (en rouge) influence sociale, pour les 4 valeurs de ρ, en fonction des types de comportement identifiés dans la Figure 5.3B. Clairement, la précision collective augmente après influence sociale, comme prévu, ce qui signifie que les individus se sont dans l'ensemble rapprochés de la vraie valeur. On remarque aussi que plus l'apport d'information est grand (ρ augmente), meilleure est la performance collective.

Il est également intéressant de voir que la catégorie de comportement dont l'amélioration est la plus grande, et qui atteint aussi la meilleure précision après influence sociale dès $\rho > 20\,\%$, est celle qui consiste à adopter l'opinion des autres. Il peut paraître trivial de prime abord que suivre des avis de meilleure qualité amène une meilleure précision des estimations. Mais il faut coupler ce résultat avec deux autres qui permettent de l'éclairer:

1. même lorsqu'aucune information n'est ajoutée ($\rho = 0\,\%$), la précision des individus qui suivent l'opinion des autres devient aussi bonne que celle des autres catégories de comportement après influence sociale;

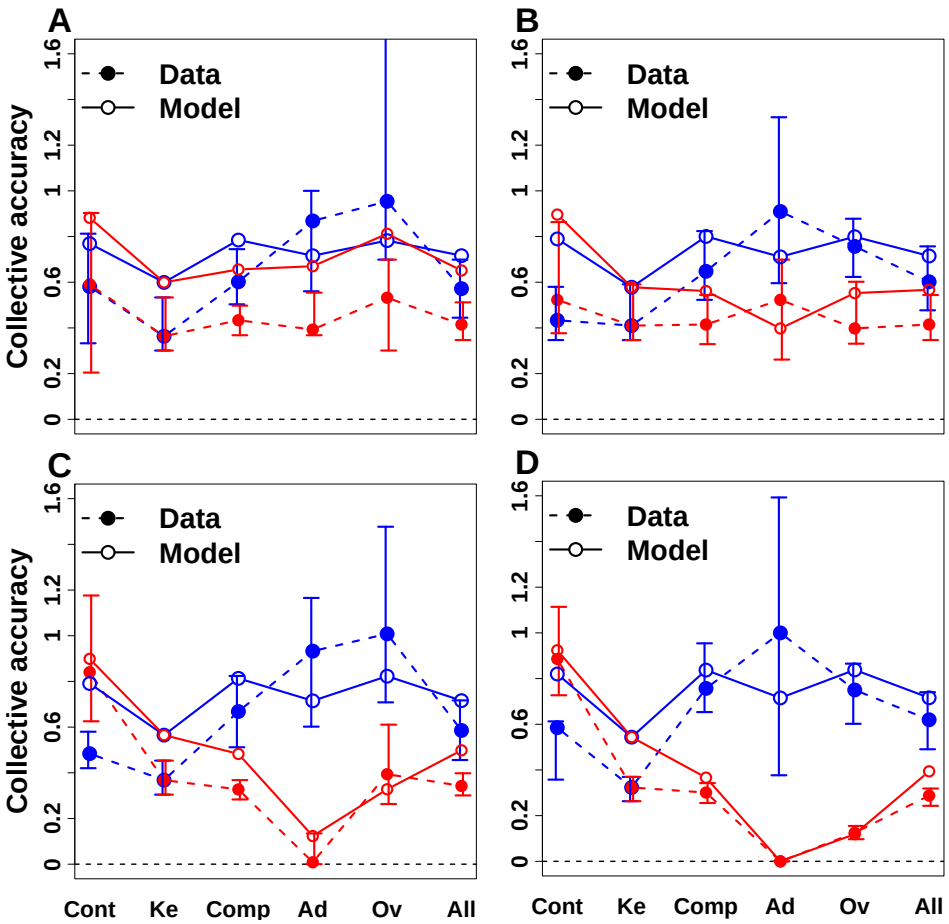

Figure 5.9: Précision collective (distance médiane des estimations individuelles à la vraie valeur) avant (bleu) et après (rouge) influence sociale, pour les 5 catégories de comportement identifiées en Figure 5.4B et pour le groupe entier (All): A. $\rho = 0\,\%$; B. $\rho = 20\,\%$; C. $\rho = 43\,\%$; D. $\rho = 80\,\%$. À mesure que ρ augmente, les individus améliorent leur précision après influence sociale. De façon intéressante, le comportement qui consiste à adopter l'information sociale mène à la meilleure amélioration de la précision, et à la meilleure précision tout court après influence sociale pour $\rho > 20\,\%$. Les cercles pleins correspondent aux données expérimentales, tandis que les cercles vides représentent les prédictions du modèle.

2. avant influence sociale, les individus qui suivent sont en moyenne (ou plutôt "en médiane") les moins précis (voir remarque ci-dessous).

Ces trois résultats mis côte à côte, nous comprenons alors que faire confiance à l'opinion des autres est la meilleure stratégie pour améliorer ses performances, que l'on soit dans un environnement clos (pas d'information extérieure, autre que celle détenue par le groupe; $\rho = 0$), ou dans un environnement où l'on peut trouver de l'information de qualité ($\rho > 0$).

Dans la section suivante, nous étudierons la suite logique, à savoir ce qu'il se passe lorsque l'environnement propose de l'information erronnée.

Remarque: avant influence sociale, les individus qui gardent leur opinion sont les plus précis, tandis que ceux qui ont tendance à suivre l'opinion des autres sont les moins précis. Deux raisons complémentaires peuvent expliquer ce phénomène:

- tout d'abord par des raisons de confiance justifiée: plus on a de connaissances sur une question, plus on est confiant dans sa réponse, et moins on a tendance à tenir compte de l'opinion des autres.

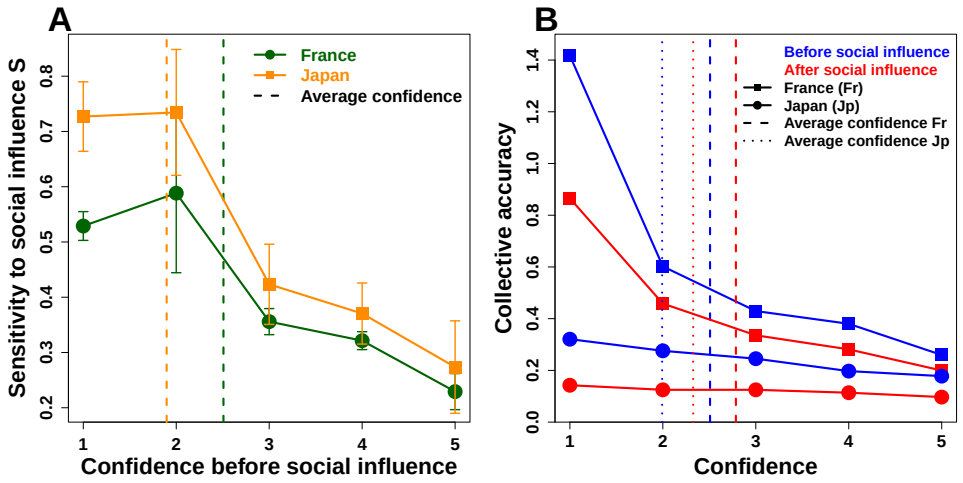

Figure 5.10: A. Sensibilité moyenne à l'influence sociale S en fonction de la confiance rapportée en France (vert) et au Japon (orange). S est négativement corrélé à la confiance, ce qui signifie que les sujets les plus confiants sont aussi les moins sensibles à l'influence sociale. La sensibilité moyenne à l'influence sociale est plus élevée au Japon qu'en France, et inversement, la confiance moyenne est plus élevée en France qu'au Japon. Cela est d'autant plus surprenant que les questions étaient plus "difficiles" en France. Ceci suggère des différences culturelles dans l'expression de la confiance et l'attention aux opinions des autres. B. Précision collective (médiane des valeurs absolues des estimations log-transformées), avant (bleu) et après influence sociale (rouge), pour les expériences réalisées en France (carrés) et au Japon (cercles). La précision collective est positivement corrélée à la confiance, ce qui signifie que les individus les plus confiants sont aussi les plus précis dans leurs réponses (réponses les plus proches de 0). Les individus dont l'estimation personnelle se rapproche de l'information sociale sont plus précis en moyenne parce que l'information sociale elle-même est en moyenne plus précise que les estimations aléatoires. La précision moyenne est meilleure au Japon qu'en France, à cause de la différence dans la difficulté des questions. La confiance moyenne augmente et la précision s'améliore après l'influence sociale.

- ensuite par une raison plus technique, venant directement de la Figure 5.7A: puisque les individus ont tendance à davantage garder leur opinion quand l'information sociale est proche de leur opinion, et que l'information sociale a davantage de chance d'être proche de la vraie valeur lorsque $\tau = 3$ (petit effet de sagesse des foules), il découle qu'il y a plus de chance de garder son opinion lorsqu'on est proche de la vraie valeur. Cet effet apparaît dans le modèle, mais est très léger, c'est pourquoi il n'est que complémentaire du premier.

Impact de la Qualité de l'Information sur la Précision des Estimations

Ici nous faisons varier la valeur T_I de l'information fournie par les agents virtuels.

La Figure 5.11 montre la performance et la précision collectives du groupe, avant (bleu) et après (rouge) influence sociale, pour $\rho = 20$ et $40\,\%$, en fonction de $\alpha = \frac{V}{\sigma}$, où V est l'information fournie par les agents virtuels (en coordonnées logarithmiques) et σ la dispersion des estimations (largeur de la distribution) pour la question considérée. α représente l'erreur, normalisée pour chaque question, de l'information fournie par les agents virtuels par rapport à la vraie valeur.

Le premier résulat surprenant est que le groupe améliore sa performance et sa précision sur de très larges gammes de α, ce qui signifie que même en présence d'information environnementale erronée, le groupe est capable d'utiliser l'information sociale pour améliorer ses performances. Un explication partielle réside dans la tendance des humains à sous-estimer les quantités, dont nous avons parlé plus haut. En effet, leur fournir une information qui surestime la vraie valeur peut compenser ce biais, et donc améliorer les performances. Mais ceci n'est pas suffisant pour expliquer la large gamme de valeurs pour lesquelles les performances du groupe s'améliorent.

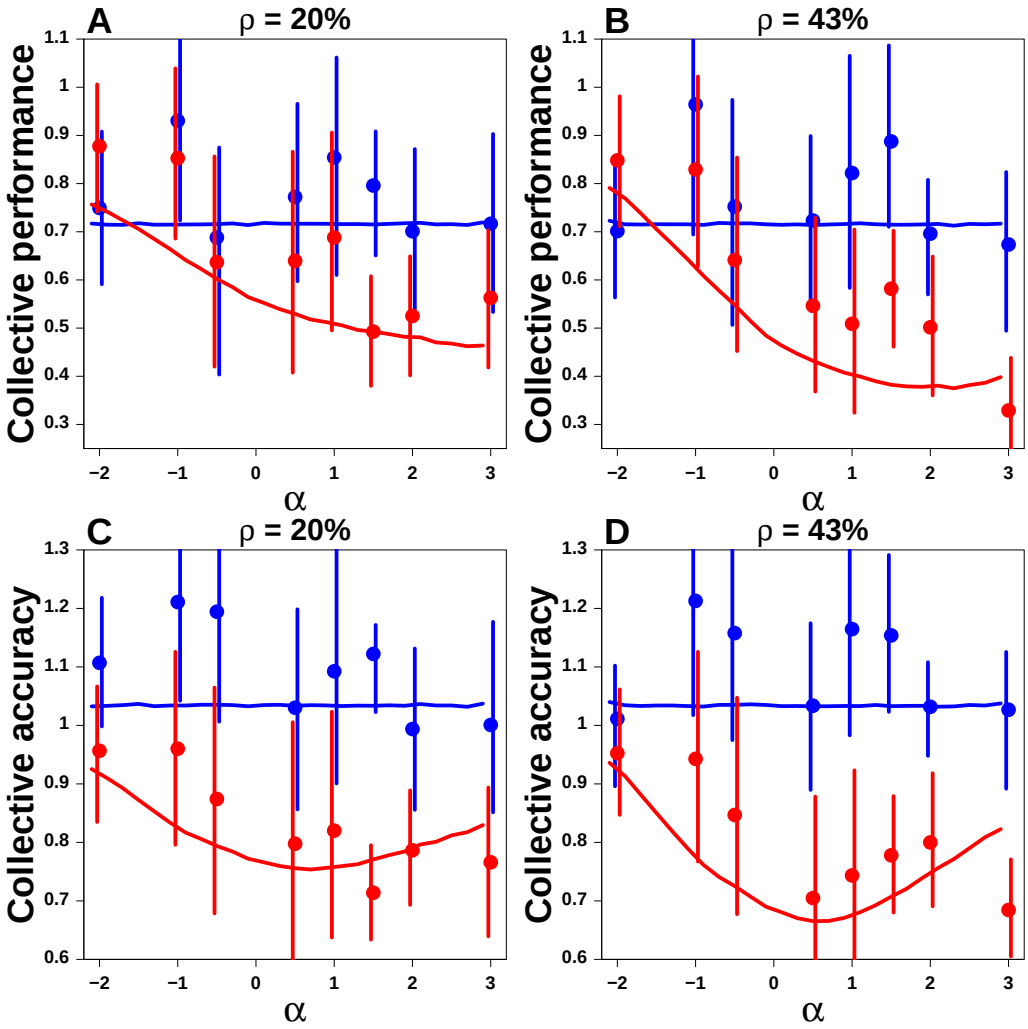

Figure 5.11: Performance (A et B) et précision (C et D) collectives en fonction de $\alpha = \frac{V}{\sigma_p}$, avant (bleu) et après (rouge) influence sociale, pour $\rho = 20\,\%$ (A and C) et $\rho = 43\,\%$ (B and D), pour le groupe complet. Les points sont les données expérimentales, et les lignes continues les simulations du modèle.

Nous avons ensuite regardé l'impact de α sur les profils de comportements, définis de manière similaire à ce que nous avons dit précédemment: en triant les réponses des individus par sensibilité à l'influence sociale S croissante, nous avons défini le profil de comportement "confiant" comme le tiers des réponses ayant le S le plus faible, le profil "suiveur" comme le tiers des réponses ayant le S le plus élevé, et le profil "moyen" comme le tiers intermédiaire.

Les Figures 5.12 et 5.13 montrent la performance et la précision collectives en fonction de α, pour ces trois catégories. Pour le comportement "confiant", comme on pouvait s'y attendre, α n'a pas plus d'impact que ρ. Ce comportement consiste à très peu tenir compte de l'avis des autres. Le profil "moyen" ressemble beaucoup au profil général, comme on l'attendait également.

Le résultat majeur vient du comportement "suiveur": la performance et la précision s'améliorent considérablement, et sur une très large gamme de valeurs de α. Autant il pouvait paraître intuitif que ce comportement amène de bonnes performances lorsque de l'information de qualité était disponible, autant cette capacité des suiveurs à utiliser de la mauvaise information pour améliorer la précision de leurs estimations est inattendue.

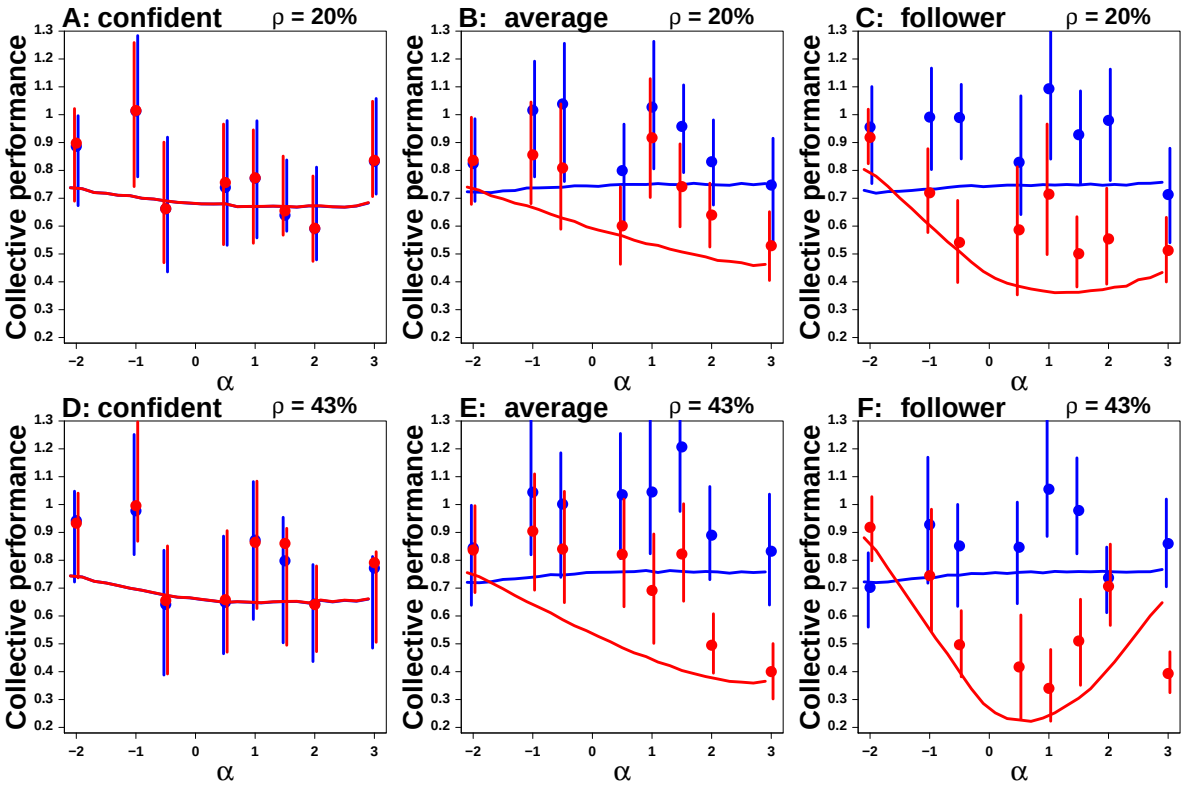

Figure 5.12: Performance collective en fonction de α, avant (bleu) et après (rouge) influence sociale, pour $\rho = 20\,\%$ (A, B et C) et $\rho = 43\,\%$ (D, E et F). A. et D. Comportement confiant; B. et E. Comportement "moyen"; C. et F. Comportement suiveur. Les points sont les données expérimentales, et les lignes continues les simulations du modèle.

Les raisons exactes de ce phénomène sont encore en cours d'étude, mais les données nous permettent d'ors et déjà d'affirmer que faire confiance aux autres est la meilleure stratégie pour améliorer ses performances, même en présence d'information fausse. Seule de l'information extrêmement trompeuse peut faire que garder son opinion soit une meilleure stratégie.

Conclusion

Quantifier la façon dont l'information sociale affecte les estimations et les opinions individuelles est une étape cruciale vers la compréhension et la modélisation de la dynamique des choix collectifs ou de la formation des opinions [133].

Dans ce travail, nous avons mesuré et modélisé l'impact de l'information sociale à aux échelles individuelle et collective dans des tâches d'estimation *séquentielles*. Plus précisément, nous contrôlions rigoureusement l'information délivrée aux sujets au moyen d'agents virtuels insérés dans la séquence des estimations – à l'insu des sujets – et fournissant une valeur de notre choix. La fraction ρ de ces agents définissait la *quantité* d'information fournie, et la valeur fournie sa *qualité*.

Ces agents virtuels peuvent être vus soit comme une source externe d'information accessible aux individus (par exemple Internet, les réseaux sociaux, les médias ...), soit comme un groupe très cohésifs d'individus ayant tous la même opinion et très confiants dans cette opinion (ayant tous $S = 0$), comme cela peut arriver avec la "pensée de groupe" [134].

Les distributions des estimations dans les quatre expériences que nous avons menées étaient plus proches de distributions de Laplace que de distributions gaussiennes, contrairement à ce

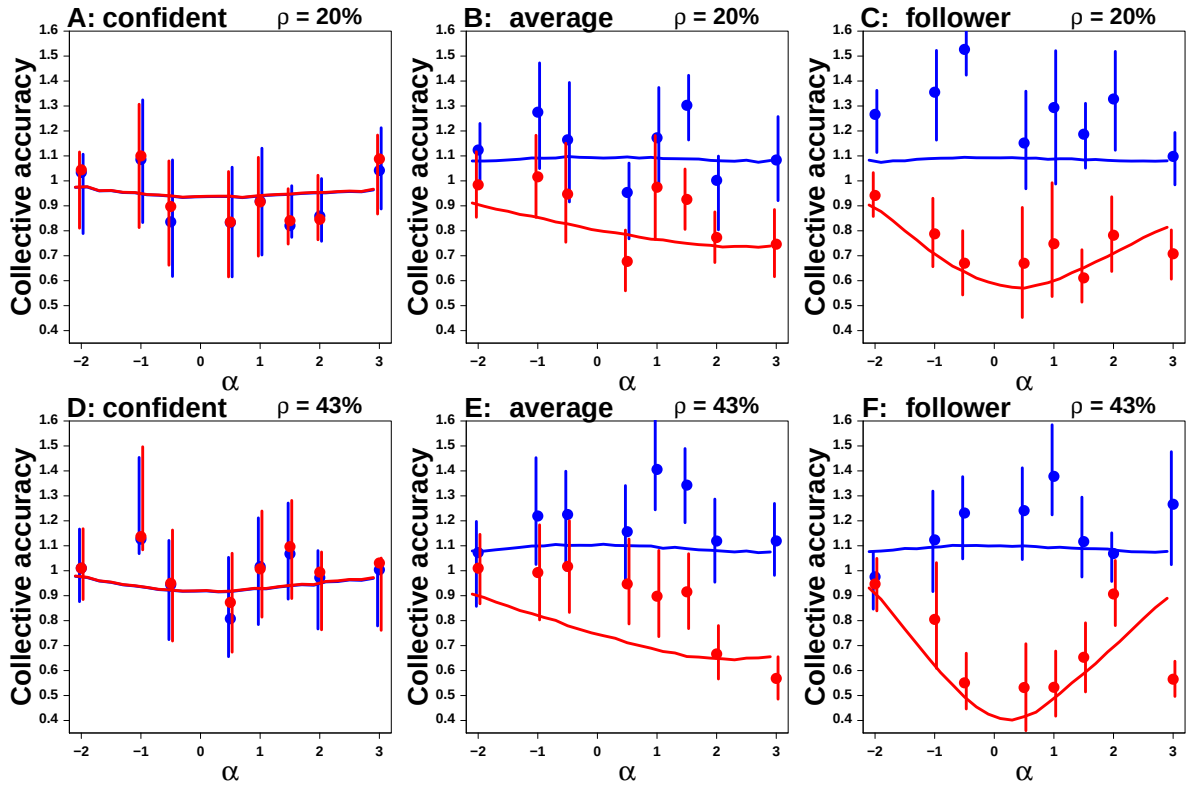

Figure 5.13: Précision collective en fonction de α, avant (bleu) et après (rouge) influence sociale, pour $\rho = 20\,\%$ (A, B et C) et $\rho = 43\,\%$ (D, E et F). A. et D. Comportement confiant; B. et E. Comportement "moyen"; C. et F. Comportement suiveur. Les points sont les données expérimentales, et les lignes continues les simulations du modèle.

qui est généralement suggéré dans la littérature. Les distributions Gaussienne et de Laplace appartiennent à la famille des Distributions Normales Généralisées, et la proximité plus ou moins grande des distributions expérimentales de l'une ou l'autre de ces distributions dépend de la connaissance préalable détenue par les individus d'un groupe sur une quantité à estimer: quand cette connaissance préalable est faible (respectivement forte), la distribution des estimations log-transformées est proche d'une distribution de Laplace (respectivement Gaussienne).

Nous avons montré qu'après influence sociale, le centre des distributions d'estimations se rapproche de la vraie valeur et sa largeur diminue, ce qui se traduit par une amélioration de la performance et de la précision collectives. Cette amélioration augmente avec ρ, à savoir avec la quantité d'information fournie aux individus.

La distribution des sensibilités à l'influence sociale S est en forme de cloche avec deux pics à $S = 0$ (garder son opinion) et $S = 1$ (adopter l'information sociale), ce qui conduit à la définition de traits sociaux robustes (individus confiants, suiveurs et "moyens"). Nos résultats ont également révélé que la sensibilité à l'influence sociale des sujets augmente linéairement avec la différence entre leur estimation personnelle et l'information sociale qu'ils reçoivent.

De façon intéressante, nous avons constaté que les individus qui ont tendance à *adopter* l'information sociale, bien qu'ils soient les moins précis avant influence sociale, deviennent au moins aussi précis que les autres après, et atteignent une précision presque parfaite dès que $\rho > 20\,\%$. Ceci suggère que donner du crédit à l'information sociale est la meilleure stratégie pour améliorer ses performances, au moins dans un environnement clos (dont on ne peut tirer aucune information) ou lorsque l'environnement fournit de l'information fiable.

Nous avons ensuite étudié l'impact d'informations incorrectes sur la performance et la précision collectives, pour l'ensemble du groupe ainsi que pour les sous-groupes de sujets confiants, "moyens" et suiveurs. Nous avons constaté que fournir des informations incorrectes peut aider un groupe à de meilleures performances que pas du tout d'information, et même meilleures qu'en fournissant la vraie valeur elle-même, par compensation de la tendance humaine à sous-estimer les quantités.

Le motif de cet effet intéressant n'est pas trivial et révèle des effets non linéaires. En particulier, l'impact d'une information incorrecte n'est pas symétrique: la performance collective peut être améliorée en fournissant des informations incorrectes surestimant la vérité jusqu'à plusieurs ordres de grandeur, alors qu'elle se dégrade beaucoup plus vite si l'information délivrée sous-estime la vérité.

Enfin, de façon très contre-intuitive, nous avons montré que les suiveurs parvenaient à de meilleures performances après influence sociale, sur une large gamme de valeurs incorrectes, suggérant que faire confiance aux autres est la meilleure stratégie pour améliorer sa précision, non seulement quand l'information disponible est parfaitement exacte, mais aussi dans des environnements incertains où l'information disponible est (parfois très) trompeuse.

Cette conclusion est renforcée par le fait que les personnes confiantes sont généralement meilleures que les autres avant l'influence sociale, mais qu'elles sont les moins performantes après l'influence sociale, même lorsque les informations fournies sont très éloignées de la vérité - les mécanismes cognitifs et comportementaux précis qui sous-tendent ces résultats fascinants sont toujours à l'étude.

Dans l'ensemble, nous avons constaté que les individus, même lorsqu'ils ont très peu de connaissances sur une quantité à estimer, parviennent à utiliser la connaissance de leurs pairs ou de l'information provenant de l'environnement pour améliorer leur précision individuelle et collective, tant que cette information environnementale n'est pas trop éloignée de la vérité.

En fin de compte, mieux comprendre ces processus ouvrira de nouvelles perspectives pour développer des systèmes d'information visant à renforcer la coopération et la collaboration dans les groupes humains, aidant ainsi les foules à devenir plus intelligentes [135, 136].

Impact de la Quantité d'Information Filtrée sur les Processus de "Séparation de Phase" chez les Humains

De façon assez similaire au projet précédent, ce projet a mêlé expériences, modélisation théorique et analyse de données, qui se sont entre-nourries les unes les autres.

Introduction et Dispositif Expérimental

Dans de nombreuses sociétés animales, les individus peuvent s'auto-organiser collectivement pour accomplir des tâches très complexes, comme chercher de la nourriture [137], chasser [138], construire des nids [139], éviter les prédateurs [140, 90, 141], etc. Ces phénomènes complexes se produisent sans contrôle externe et émergent plutôt des interactions locales entre individus [70]. En raison du caractère intrinsèquement limité de leurs capacités de calcul et de l'information disponible sur leur environnement (*rationalité bornée*), les réponses comportementales chez de nombreux animaux sont souvent déclenchées localement par des informations provenant de voisins proches [87, 88, 89, 90, 91, 92].

Des phénomènes similaires ont été observés dans les sociétés humaines, comme dans le trafic automobile [143], l'organisation urbaine [144] ou les mouvements collectifs de piétons [102]. Les phénomènes collectifs tels que la formation des voies ou des sentiers [102, 145, 146, 147], ou les dynamiques de foule [95, 148, 149, 150] émergent des interactions locales entre individus [66]. Les piétons sont donc un terreau très fertile pour l'étude des comportements collectifs, et il reste beaucoup à comprendre avant de pouvoir, par exemple, réduire drastiquement les désastres survenant lors de paniques de foule. Des solutions partielles ont été proposées, telles que l'amélioration des systèmes d'évacuation ou la conception d'infrastructures pour faciliter le flux au niveau des goulets d'étranglement ou des carrefours. Nous pensons qu'une approche complémentaire basée sur l'interaction des piétons avec des *systèmes de traitement de l'information* peut être fructueuse. Cependant, la façon dont de tels systèmes pourraient interagir avec le comportement collectif des piétons (voire le contrôler) est encore largement inconnue.

Pour aborder ce problème, nous avons conçu des tâches spécifiques dans lesquelles les piétons devaient s'appuyer sur un tel système: nous avons assigné une "couleur" (sous-groupe) au hasard à des sujets dans des groupes de 22 piétons placés dans une arène circulaire. Ils ne connaissaient ni leur propre couleur ni celle des autres membres du groupe (pas d'indice visuel). La tâche consistait à se séparer en groupes de la même couleur (*séparation de phase*).

Pour aider les piétons à accomplir leurs tâches, nous avons conçu un système imitant un *dispositif sensoriel* (tel la rétine), capable de condenser l'information accessible (couleur et position de tous les piétons dans l'arène) en un bit d'information: des tags attachées aux épaules des sujets transmettaient entièrement les positions et couleurs de tous les piétons en temps réel à un serveur central, et délivraient en retour un signal acoustique (un "bip") sous des conditions bien spécifiques:

- *Règle "majoritaire"*: le tag d'un sujet émet un bip si la majorité de ses k plus proches voisins ($k = 1, 3, 5, 7, 9, 11, 13$) est d'une couleur différente de la sienne.
- *Règle "exclusive"*: le tag d'un sujet émet un bip si *au moins un* de ses k plus proches voisins ($k = 1, 2, 3, 4$) est d'une couleur différente de la sienne.
- *Règle "décalée"*: le tag d'un sujet émet un bip si la majorité de ses k, $(k+1)$ et $(k+2)$-ème plus proches voisins ($k = 1, 2, 3, 4$) est d'une couleur différente de la sienne.

Pour accomplir la tâche, les sujets n'avaient accès qu'à ces signaux acoustiques et n'étaient pas au courant de la règle précise: nous leur avions simplement dit qu'ils biperaient chaque

fois que leur "environnement" serait de l'autre couleur que la leur. En faisant varier k dans la condition *majoritaire*, nous contrôlions le nombre de voisins considérés dans le calcul du bip, et donc la *quantité* d'information traitée. En variant k dans les conditions *exclusive* et *décalée*, nous avons manipulé la complexité de la tâche.

Nous avons défini la performance du groupe comme le temps de ségrégation (plus le temps est court mieux c'est) et, par analogie avec les phénomènes de séparation de phase, le nombre de clusters à l'instant final (moins il y a de clusters mieux c'est). Nous avons examiné l'impact de k sur ces deux mesures dans toutes les conditions, et avons trouvé une dépendance en forme de U du temps de ségrégation par rapport à k, alors que le nombre moyen de clusters à l'instant final diminue avec k, et sature à des valeurs de k qui varient avec la taille du groupe.

Pour mieux comprendre les mécanismes qui sous-tendent ces résultats, nous avons construit et calibré un modèle de déplacement de piétons basé sur l'approche dite des *forces sociales* [99]. Cette approche a réussi à bien décrire qualitativement certains phénomènes collectifs [95, 148, 150], mais les forces d'interaction ont généralement été simplement fittées aux données expérimentales [151, 152]. Plus tard, des efforts ont été faits pour *extraire* la forme des interactions à partir des données et ainsi améliorer la précision des modèles de force sociale [103]. Nous poursuivons dans cette direction et proposons d'aller plus loin dans la qualité de la description des interactions, en combinant des méthodes récentes qui ont fait leurs preuves dans la mesure d'interactions chez les poissons [153, 107].

L'approche générale consiste à collecter des données de mouvement pour des individus isolés (mouvements spontanés et interactions avec l'environnement physique), puis pour deux individus seuls (interactions entre individus) et enfin pour des groupes. À partir des données avec un et deux individus, on peut définir un modèle, les prédictions pour le mouvement collectif duquel sont ensuite comparées aux données obtenues à partir du groupe entier [153]. Dans nos expériences, nous avons demandé à des groupes de 1, 2 ou 22 piétons de se déplacer de façon aussi aléatoire que possible dans des arènes circulaires (on a demandé à des groupes de 22 personnes de le faire pendant 45 secondes avant le début des bips). Ensuite, pour construire le modèle, nous avons utilisé une procédure permettant de reconstruire la forme fonctionnelle des interactions directement à partir des données [107].

Bien que l'étude des forces d'interaction entre piétons ne soit pas encore complète, la version actuelle de notre modèle est déjà capable de reproduire très précisément des mesures de mouvement collectif aussi fines que l'autocorrélation des vitesses, le processus de diffusion (bornée) et les distributions de vitesses, positions, plus proches voisins et angles au mur. Nous avons reproduit avec succès les résultats observés dans la phase de ségrégation, et nous avons pu faire des prédictions pour les processus de ségrégation à différentes tailles de groupes.

Modèle et marche aléatoire

Suivant des travaux antérieurs sur les dynamiques de piétons, nous modélisons les piétons par des particules suivant un mouvement Brownien (marche aléatoire) et soumis à des *forces sociales* qui représentent leur "motivation interne" de se déplacer [99, 100, 98]:

$$\frac{d\vec{r}_i(t)}{dt} = \vec{v}_i(t), \tag{5.44}$$

$$\frac{d\vec{v}_i(t)}{dt} = -A(\vec{v}_i(t)) + \vec{\eta}_i(t) + \vec{F_{\mathrm{w}_i}}(t) + \sum_{j=1,\, j\neq i}^{N} \vec{F_{\mathrm{h}_{ij}}}(t), \tag{5.45}$$

où i est l'indice désignant les piétons, $\vec{r_i}$ et $\vec{v_i}$ sont respectivement la position (par rapport au centre de l'arène) et la vitesse de l'individu i. A est l'auto-propulsion, $\vec{\eta}$ un bruit blanc Gaussien, et $\vec{F}_{\mathrm{w}}$ et $\vec{F}_{\mathrm{h}}$ sont respectivement les forces d'interaction du piéton i avec le mur et avec les autres piétons j.

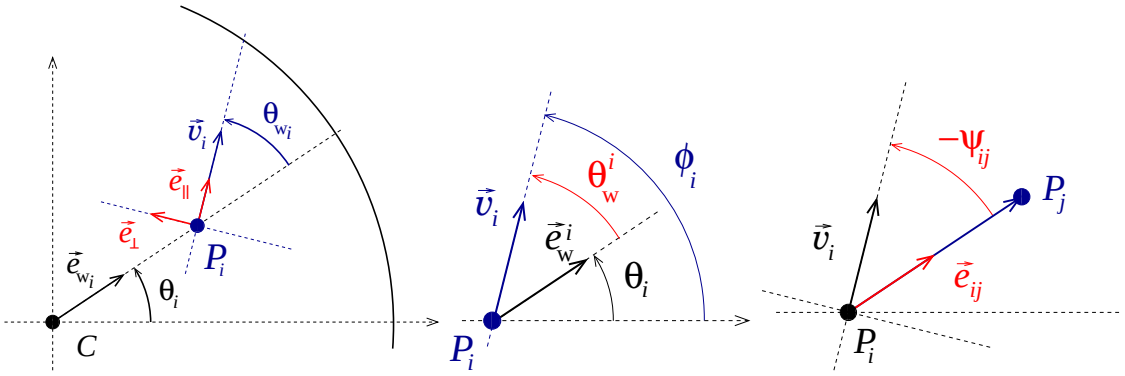

Figure 5.14: Notations utilisées pour caractériser la position et le mouvement des piétons: le piéton P_i marche dans un cercle centré en C, avec une vitesse $\vec{v}_i$ et un angle au mur θ_{w_i}. $\vec{e}_{\mathrm{w}_i}$ est le vecteur radial unitaire, $\vec{e}_{\parallel}$ le vecteur unitaire dans la direction du mouvement, $\vec{e}_{\perp}$ le vecteur unitaire perpendiculaire à la direction du mouvement, et $\vec{e}_{ij}$ le vecteur unitaire dans la direction $P_i\vec{P}_j$ (P_j est un autre piéton). θ_i est l'angle entre $\vec{e}_{\mathrm{w}_i}$ et l'horizontale, ϕ_i l'angle entre la direction du mouvement et l'horizontale, et Ψ_{ij} l'angle entre $\vec{v}_i$ et $\vec{e}_{ij}$.

Nous avons introduit une méthode qui permet d'extraire directement la forme fonctionnelle des interactions à partir des données elles-mêmes. L'application de cette méthode au cas à 2 piétons est toujours en cours, mais nous avons pu déterminer avec précision les formes de A et F_{w}:

$$A(\vec{v}) = \frac{v - \bar{v}}{\tau_0}\vec{e}_{\parallel} \tag{5.46}$$

est l'accélération des piétons vers leur vitesse de confort $\bar{v}$. Quant à l'interaction avec le mur, on s'attend à ce qu'elle dépende de la vitesse et de l'angle d'incidence d'un piéton vers le mur, ainsi que de la distance entre le piéton et le mur:

$$\vec{F}_{\mathrm{w}}(r_{\mathrm{w}}, \theta_{\mathrm{w}}, v) = -B_{\mathrm{w}}(v) f_{\mathrm{w}}(r_{\mathrm{w}}) g_{\mathrm{w}}(\theta_{\mathrm{w}})\vec{e}_{\mathrm{w}}, \tag{5.47}$$

où $\vec{e}_{\mathrm{w}}$ est le vecteur radial unitaire (voir Figure 5.14). L'analyse a montré que l'on pouvait négliger la dépendance en la vitesse ($B_{\mathrm{w}}(v) = 1$), et que les fonctions f_{w} et g_{w} ont la forme suivante:

$$g_{\mathrm{w}}(\theta_{\mathrm{w}}) = a_{\mathrm{w}_0} + a_{\mathrm{w}_1}\cos(\theta_{\mathrm{w}}) + a_{\mathrm{w}_2}\cos(2\theta_{\mathrm{w}}) + a_{\mathrm{w}_3}\cos(3\theta_{\mathrm{w}}) + a_{\mathrm{w}_4}\cos(4\theta_{\mathrm{w}}) \tag{5.48}$$

$$f_{\mathrm{w}}(r_{\mathrm{w}}) = \begin{cases} a_{\mathrm{w}}\left(e^{\frac{-r_{\mathrm{w}}}{l_{\mathrm{w}}}} - e^{\frac{-r_{\mathrm{w_c}}}{l_{\mathrm{w}}}}\right) & \text{if } r_{\mathrm{w}} < r_{\mathrm{w_c}} \\ 0 & \text{otherwise} \end{cases} \tag{5.49}$$

où r_{w} est la distance au mur, $r_{\mathrm{w_c}}$ est la distance critique au mur au-delà de laquelle le mur n'a plus d'effet sur les piétons, et l_{w} est la portée typique de l'interaction.

Pour l'interaction entre piétons, nous avons utilisé une fonction similaire (en attendant d'extraire la véritable forme de l'interaction à partir des données), mais un peu plus répulsive à courte portée:

$$f_{\mathrm{h}}(r_{ij}) = \begin{cases} a_{\mathrm{h}}\left(e^{-\left(\frac{r_{ij}}{l_{\mathrm{h}}}\right)^2} - e^{-\left(\frac{r_{\mathrm{h_c}}}{l_{\mathrm{h}}}\right)^2}\right) & \text{if } r_{ij} < r_{\mathrm{h_c}} \\ 0 & \text{otherwise} \end{cases} \tag{5.50}$$

où r_{ij} est la distance entre le piéton i et le piéton j, $r_{\mathrm{h_c}}$ est a distance inter-individuelle critique au-delà de laquelle l'interaction disparaît, et l_{h} est la portée typique de l'interaction.

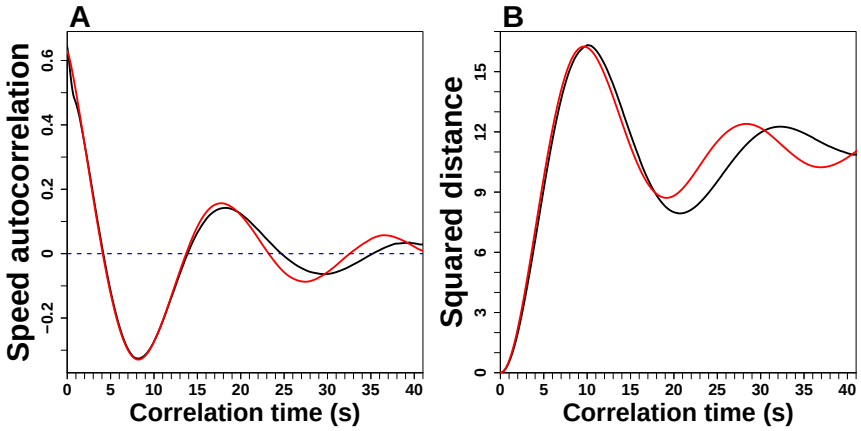

Figure 5.15: A. Fonction d'autocorrelation de la vitesse $C(t) = \langle \vec{v}_{e,i}(t'+t) \cdot \vec{v}_{e,i}(t') \rangle_{e,i,t'}$; B. Distance carrée moyenne (diffusion) $\chi(t) = \langle (\vec{r}_{e,i}(t'+t) - \vec{r}_{e,i}(t'))^2 \rangle_{e,i,t'}$, en fonction du temps de corrélation. Les données sont en noir et la simulation du modèle en rouge.

Les Figures 5.15 et 5.16 montrent la précision du modèle, qui est capable de reproduire les données expérimentales de façon très fine. Nous avons ensuite utilisé ce modèle, une fois bien calibré, pour analyser l'impact de k sur la dynamique de ségrégation.

Séparation de Phase avec Information Filtrée et Contrôlée

Durant la phase de ségrégation, les individus bipent lorsque la majorité de leurs k plus proches voisins appartiennent au sous-groupe différent du leur (pour rappel il n'y avait toujours que deux sous-groupes dans nos expériences). Nous conservons notre modèle, à ceci près que nous ajoutons un "interrupteur" ϵ_i, qui vaut 1 lorsque l'individu i bipe, et 0 sinon. Lorsque l'individu bipe, il se comporte comme lors de la marche aléatoire, et lorsqu'il s'arrête de biper, le bruit disparaît, et l'accélération vers la vitesse de confort est remplacée par une force de "friction", représentant la volonté du piéton de s'arrêter (ce qui n'est pas instantané). L'équation du mouvement s'écrit donc:

$$\frac{d\vec{v}_i(t)}{dt} = (1-\varepsilon_i(t)) \left(-\frac{\vec{v}_i(t)}{\tau}\right) + \varepsilon_i(t)\ (-A(\vec{v}_i(t)) + \vec{\eta}_i(t)) + \vec{F_{\mathrm{w}_i}}(t) + \sum_{j=1,\, j\neq i}^{N} \vec{F_{\mathrm{h}_{ij}}}(t), \qquad (5.51)$$

Dynamique Générale

Les Figures 5.17 et 5.18 montrent l'évolution de la proportion de bips et du nombre moyen de clusters en fonction du temps, pour les différentes valeurs de k.[7]

Impact de la Quantité d'Information sur les Processus de Séparation de Phase

Pour mieux visualiser l'impact de k sur la dynamique de ségrégation, nous quantifions le temps et la qualité de la ségrégation, définie par analogie avec la séparation de phase: moins il y a de clusters à la fin de la ségrégation, plus la qualité de la ségrégation est grande. Ainsi, la

[7]Pour des raisons non-encore parfaitement élucidées, les temps de ségrégation dans les expériences effectuées en Septembre 2015 étaient environ 2 fois plus courts que ceux effectués en Juin 2016. Comme le cas $k = 13$ n'a été fait qu'en Juin 2016, nous ne le présentons pas dans les figures impliquant le temps de ségrégation. Cependant, le processus de clustering n'a pas été affecté, de sorte que nous pouvions combiner toutes les données pour les graphiques correspondants.

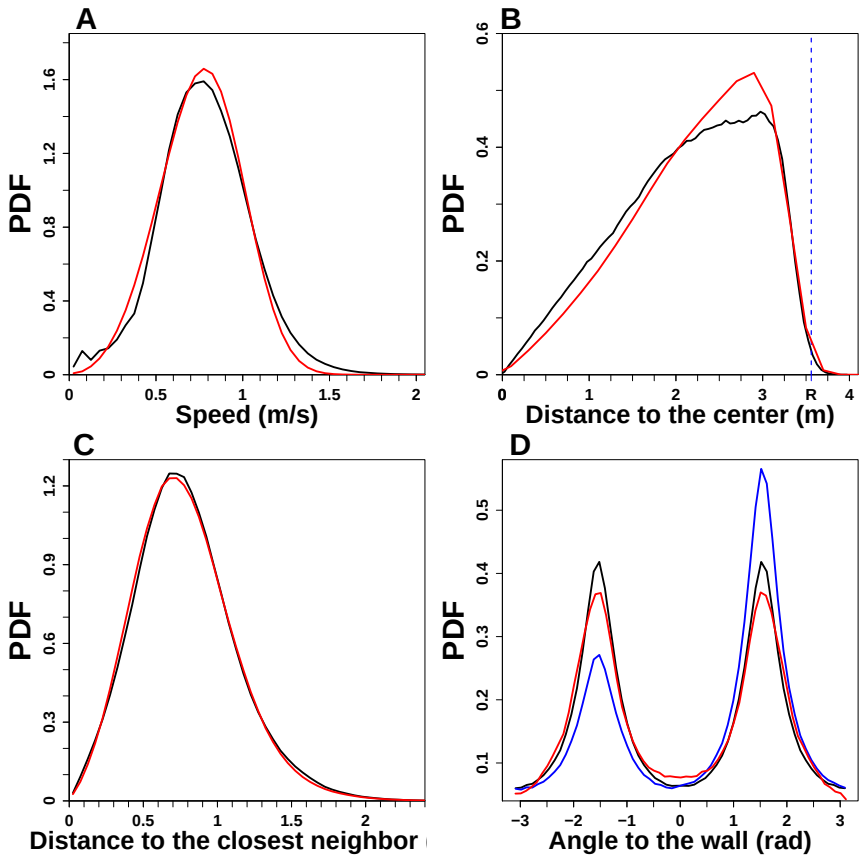

Figure 5.16: A. Distribution de vitesses; B. Distribution de positions (distances au centre); C. Distribution de distances au plus proche voisin; Les données sont en noir et les simulations du modèle en rouge. D. Distribution des angles au mur. Les données originelles sont en bleu, et ne sont pas symmétriques: les piétons préfèrent avoir le mur à leur droite. En noir sont montrées les données symmétrisées (pour chaque angle, nous considérons aussi l'angle opposé), et en rouge sont les simulations du modèle.

ségrégation est "parfaite" s'il y a deux clusters bien distincts à la fin du processus (les "phases" sont séparées). Nous définissons les trois temps de ségrégation suivants:

- $\langle t_{\mathrm{b}_{i,e}} \rangle_{i,e}$, où $t_{\mathrm{b}_{i,e}}$ est le temps total où l'individu i dans l'expérience e a bipé,
- $\langle t_{\mathrm{f}_{i,e}} \rangle_{i,e}$, où $t_{\mathrm{f}_{i,e}}$ est le dernier temps où l'individu i dans l'expérience e a bipé ($t_{\mathrm{f}_{i,e}} \geq t_{\mathrm{b}_{i,e}}$),
- $\langle t_{\mathrm{end}_e} \rangle_e$, où $t_{\mathrm{end}_e} = \mathrm{Max}_i(t_{\mathrm{f}_{i,e}})$ est le dernier temps où un individu a bipé dans l'expérience e, tous individus confondus.

La Figure 5.19 montre ces temps de ségrégation en fonction de k. L'impact de k est assez faible, bien que les Figures B et C suggèrent un optimum autour de $k = 7 \sim 9$.

La Figure 5.20 montre différents indicateurs de la qualité du clustering à la fin de la ségrégation. Le même pattern apparaît dans tous les graphes: la qualité du clustering est minimale à $k = 1$, puis augmente jusqu'à atteindre un plateau autour de $k = 7 \sim 9$. Nous avons ensuite étudié comment cette valeur évolue avec différentes tailles de groupes.

Les Figures 5.21 et 5.22 montrent les simulations du modèle pour la dépendance en k du temps de ségrégation et de la qualité du clustering (on ne montre ici que la taille moyenne des clusters et la proportion d'expériences où la ségrégation a été parfaite, car elles sont corrélées aux autres) pour des groupes de 10, 22, 30 et 42 piétons.

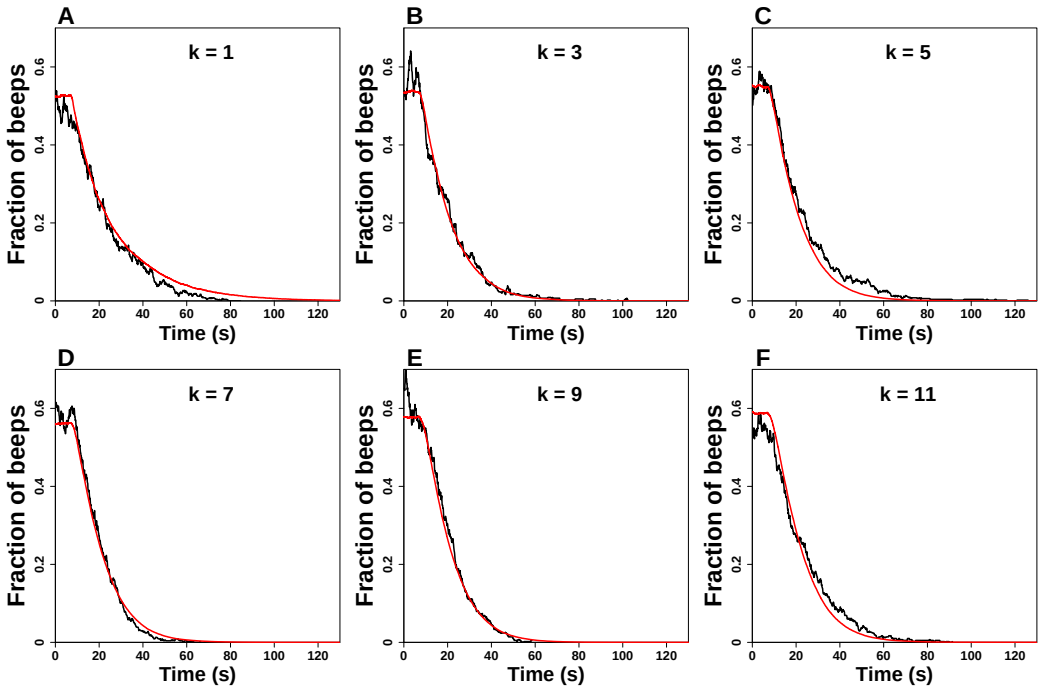

Figure 5.17: Évolution temporelle de la proportion de sujets qui bipent, pour chaque valeur de k. Les données sont en noir, et les simulations du modèle en rouge.

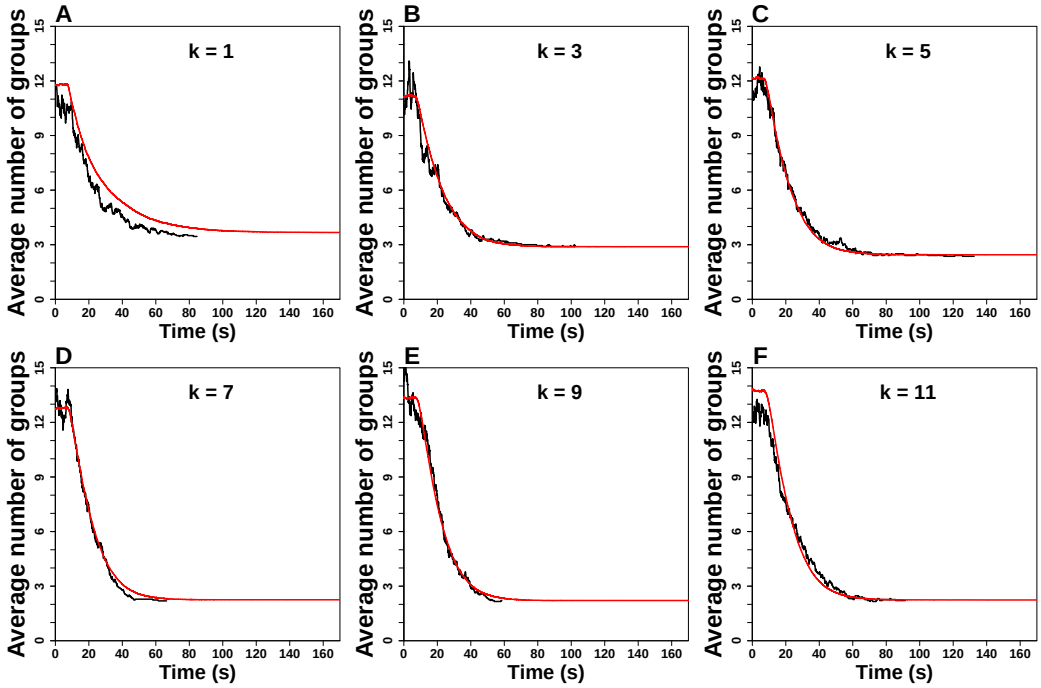

Figure 5.18: Évolution temporelle du nombre moyen de clusters, pour chaque valeur de k. Les données sont en noir, et les simulations du modèle en rouge.

Fait intéressant, la valeur optimale du temps de ségrégation (voir Figure 5.21) est indépendante de la taille du groupe: $k = 3$ pour $\langle t_b \rangle$ et $\langle t_{\text{end}} \rangle$, et $k = 5 \sim 7$ pour t_f. Globalement, sauf pour le cas $k = 1$ qui est trop bruité, les petites valeurs de k doivent être préférées pour minimiser le temps de ségrégation.

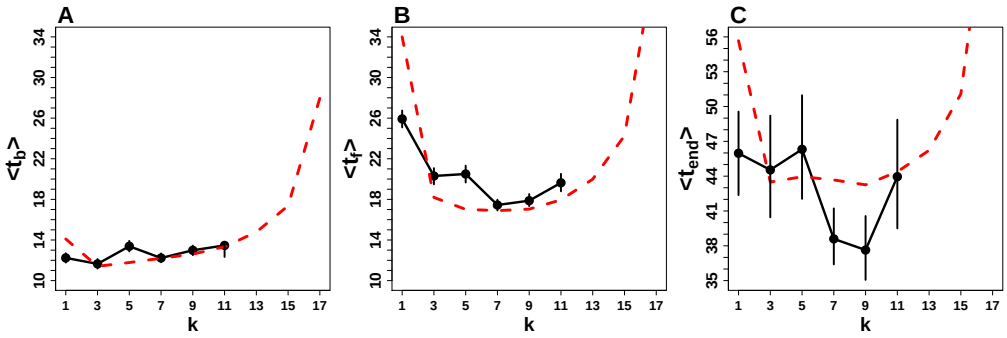

Figure 5.19: Moyenne du temps de ségrégation, défini comme: A. t_b: le temps total où un individu a bipé; B. t_f: le dernier instant où un individu a bipé; C. t_{end}: la valeur maximale des t_f pour une expérience donnée, en fonction du paramètre d'information k. Le temps de ségrégation augmente rapidement pour $k > 11$.

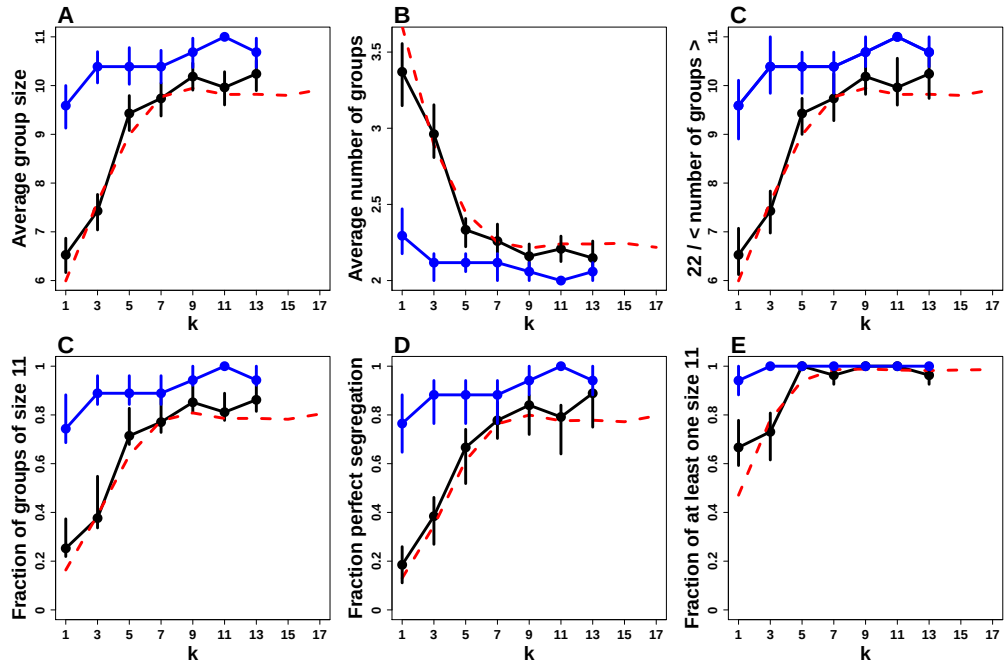

Figure 5.20: A. Taille moyenne des clusters; B. Nombre moyen de clusters; C. 22 / Nombre moyen de clusters; D. Proportion de clusters de taille 11; E. Proportion d'expériences où la ségrégation est parfaite (2 clusters de taille 11); F. Proportion d'expériences avec au moins un cluster de taille 11 à l'instant final, en fonction du paramètre d'information k. Les données sont en noir et les simulations du modèle en rouge. En bleu sont les données expérimentales lorsque la condition supplémentaire de se ségréger en deux clusters était demandée aux sujets.

En ce qui concerne le clustering (voir Figure 5.22), bien que les patterns soient globalement les mêmes pour toutes les tailles de groupes, on observe des différences notables:

- Pour $N = 10$, les piétons se séparent presque toujours parfaitement pour $k = 1$ (deux clusters de taille 5), et toujours parfaitement pour $k \geq 3$.
- Pour $N \geq 22$, le plateau (clustering optimal) est atteint pour des valeurs de plus en plus petites de k par rapport à la moitié de la taille du groupe: $k = 7 \sim 9$ pour $\frac{N}{2} = 11$, $k = 9 \sim 11$ pour $\frac{N}{2} = 15$ et $k \sim 13$ pour $\frac{N}{2} = 21$.
- De même, la taille moyenne des clusters à l'instant final ne change pas en proportion de la taille du groupe. Alors qu'il est égal à 5 pour $\frac{N}{2} = 5$ (ségrégation parfaite), il est d'environ 10 pour $\frac{N}{2} = 11$, d'environ $12 \sim 13$ pour $\frac{N}{2} = 15$ et d'environ 16 pour $\frac{N}{2} = 21$.

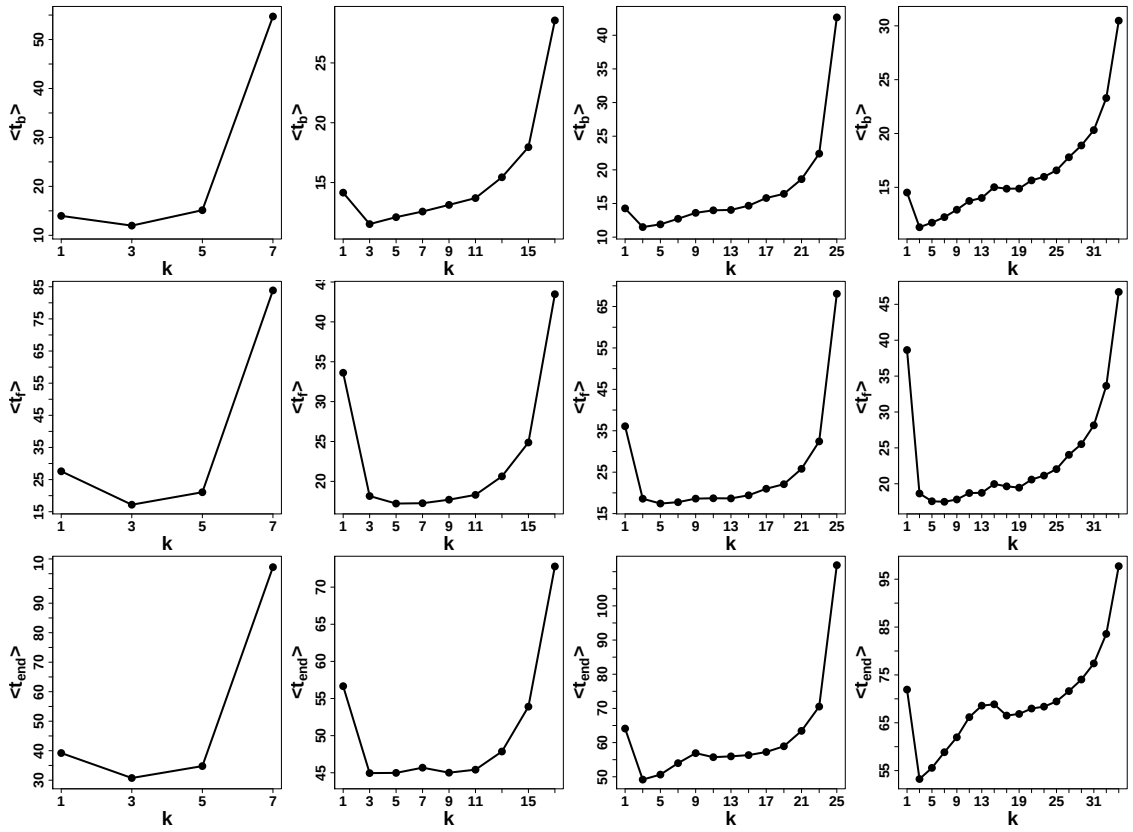

Figure 5.21: Simulations du modèle pour le temps de ségrégation défini comme t_{b}: le temps total passé par un individu à émettre des bips (première ligne); t_{f}: la dernière fois qu'un individu a émis un bip (deuxième ligne); t_{end}: la valeur maximale de t_{f} pour une expérience (troisième ligne), en fonction du paramètre d'information k, pour des groupes de 10 (première colonne), 22 (deuxième colonne), 30 (troisième colonne) et 42 (quatrième colonne) piétons.

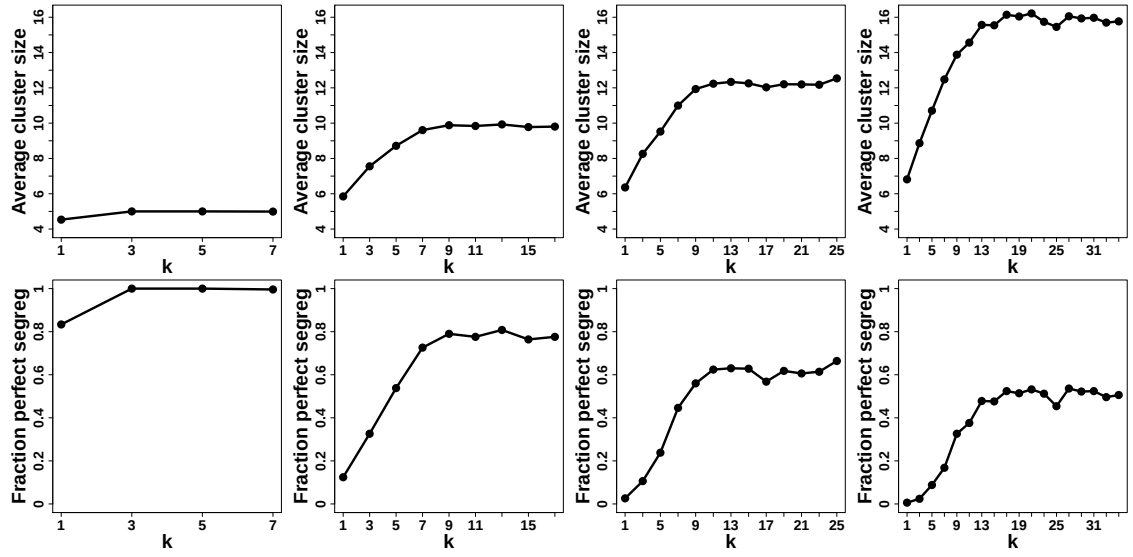

Figure 5.22: Simulations du modèle pour la taille moyenne des clusters (première ligne) et la proportion d'expériences où la ségrégation a été parfaite (deuxième ligne) en fonction du paramètre d'information k, pour des groupes de 10 (première colonne), 22 (deuxième colonne), 30 (troisième colonne) et 42 (quatrième colonne) piétons. Nous ne montrons pas les autres mesures à nouveau, car elles sont corrélées et n'apportent pas d'informations supplémentaires. Leur forme peut être déduite facilement en regardant la figure 5.20.

- De façon corrélée, la fraction d'expériences finissant par une ségrégation parfaite décroît avec la taille du groupe: 100 % pour $\frac{N}{2} = 5$, et environ 80 % pour $\frac{N}{2} = 11$, 60 % pour $\frac{N}{2} = 15$ et 50 % pour $\frac{N}{2} = 21$.

Séparation de Phase dans des Environnements plus Complexes

En plus de la condition "*majoritaire*", nous avons essayé deux autres "environnements" (conditions *exclusive* et *décalée*), pour analyser l'impact de la difficulté de la tâche sur les processus de ségrégation, et pour vérifier si notre modèle (calibré pour la condition *majoritaire*, reproduirait les résultats obtenus dans ces nouvelles conditions.

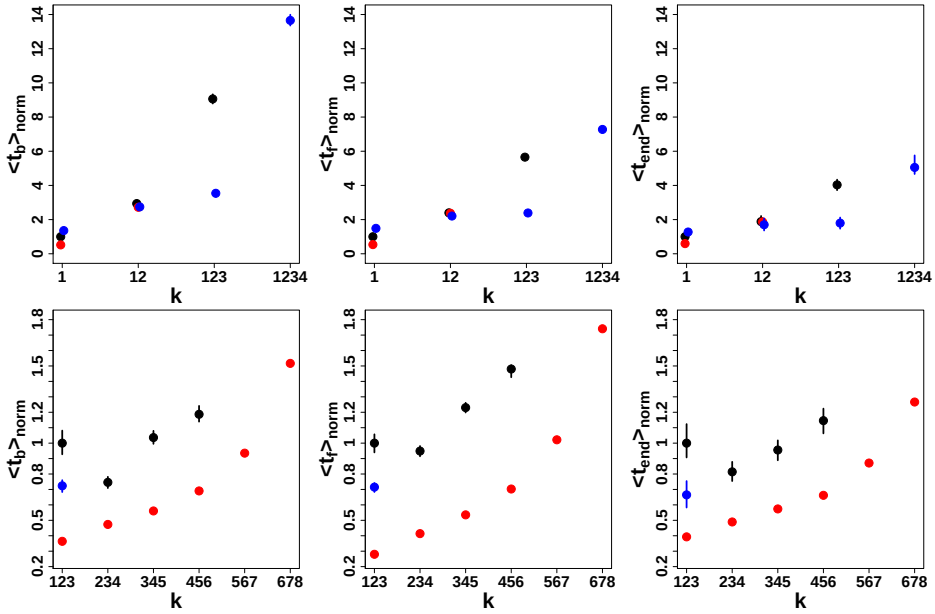

Figure 5.23: Temps de ségrégation défini comme t_b: temps total passé par un individu à émettre un bip (première colonne); t_f: la dernière fois qu'un individu a émis un bip (deuxième colonne); t_{end}: la valeur maximale de t_f pour une expérience (troisième colonne), par rapport au paramètre d'information k, dans les conditions "exclusive" (première ligne) et "décalée" (deuxième ligne). Toutes les valeurs sont normalisées par la valeur correspondant au point noir le plus à gauche (qui sert ainsi de point de référence). Les données sont en noir et les simulations du modèle sont en rouge. En bleu sont les données expérimentales pour le cas où les sujets étaient invités à se séparer en deux clusters. Notez qu'aucun point rouge n'apparaît pour les cas "123" et "1234" dans l'environnement exclusif, car les temps de ségrégation prévus par le modèle sont extrêmement longs.

Puisque les environnements *exclusif* et *décalé* n'ont été effectués qu'en juin 2016, les temps de ségrégation observés sont environ deux fois plus élevés que ceux prédits par le modèle à partir des données de septembre 2015. Ceci gardé à l'esprit, les Figures 5.23 et 5.24 montrent que notre modèle est en accord avec les données dans la plupart des cas.

Comme prévu, des tâches plus complexes entraînent des temps de ségrégation plus longs. Cependant, notre modèle prédit des temps de ségrégation beaucoup plus longs pour le cas "123" dans la condition *exclusive* que ce que montrent les données expérimentales, suggérant l'existence d'un mécanisme non encore dévoilé. Le temps de ségrégation a également été affecté, mais de façon moins spectaculaire, lorsque les sujets ont été explicitement invités à se séparer en deux clusters distincts. De façon intéressante, cette condition supplémentaire a aidé les individus à se séparer plus rapidement dans tous les cas, à l'exception de $k = 1$ dans la condition *exclusive* (ce qui équivaut à $k = 1$ dans la condition *majoritaire*).

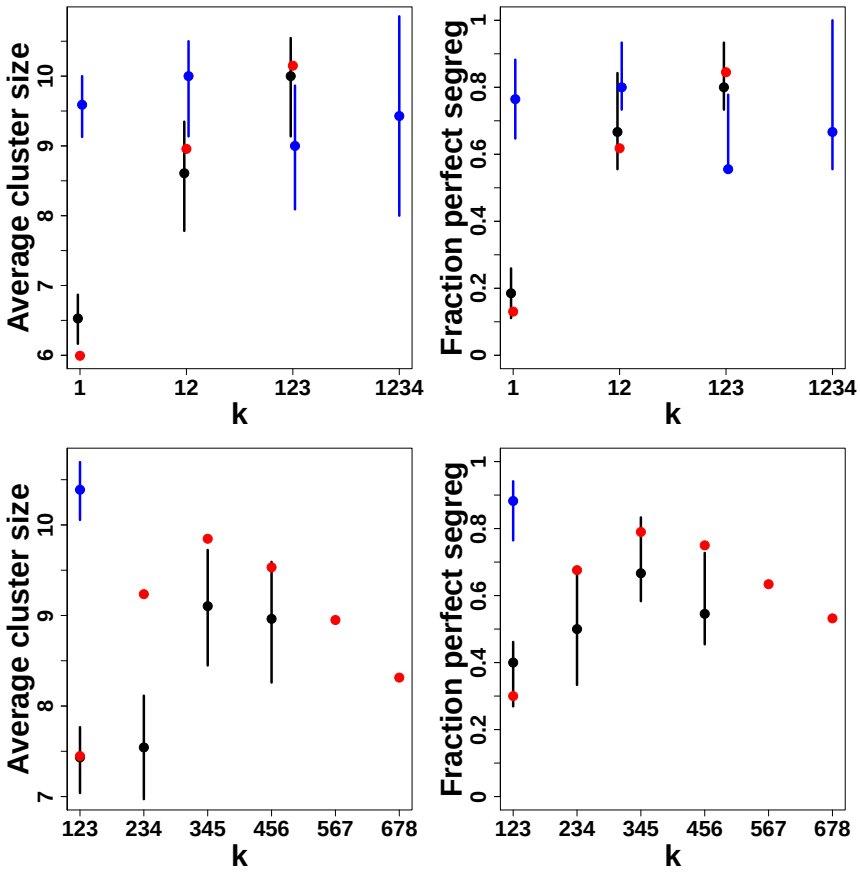

Figure 5.24: Taille moyenne des clusters (à gauche) et fraction des expériences finissant par une ségrégation parfaite (à droite) contre le paramètre d'information k, dans les conditions "exclusive" (en haut) et "décalée" (en bas). Les données sont en noir et les simulations du modèle sont en rouge. En bleu sont les données expérimentales pour la condition où les sujets ont été invités à se séparer en deux clusters bien distincts.

Comme mentionné ci-dessus, le processus de clustering n'a pas été affecté par le changement du temps de ségrégation, et les prédictions de notre modèle reproduisent bien les données. Nous trouvons que la qualité du clustering s'améliore avec le nombre de voisins considérés dans le critère *exclusif*. En effet, la taille moyenne des clusters à l'instant final augmente, tout comme la fraction des expériences dont la ségrégation est parfaite.

Nous observons aussi un phénomène très intéressant et non-intuitif dans la condition *décalée*: la qualité du clustering s'améliore si on considère non pas les trois plus proches voisins, mais les deuxième, troisième et quatrième voisins les plus proches, et encore mieux avec les troisième, quatrième et cinquième voisins les plus proches. Au-delà, la qualité du clustering commence à décroître, mais lentement et en restant meilleure que le cas classique "123" (les trois plus proches voisins).

Conclusion

Dans ce chapitre, nous avons examiné comment le comportement collectif des piétons dans les tâches de ségrégation était affecté par la quantité d'information traitée par un système de filtrage d'information. Nous avons conçu un dispositif sensoriel artificiel capable de convertir des informations complexes en entrée (couleurs et positions de tous les individus dans une arène fermée) en une simple information en sortie (un bip ou pas de bip), selon des règles spécifiques (conditions *majoritaire*, *exclusive* et *décalée*).

Les tâches consistaient à se regrouper avec des piétons de la même couleur. Pour y arriver, les sujets n'avaient accès qu'à des signaux acoustiques (bips, sortie du système de filtrage de l'information) provenant de tags électroniques attachés à leurs épaules. Ils ne connaissaient ni leur propre couleur ni celle des autres participants, et on leur demandait seulement d'essayer de ne pas biper.

Nous avons défini la qualité d'un processus de ségrégation comme étant d'autant meilleure que le temps de ségrégation est court et le nombre de groupes à l'instant final (par analogie avec la séparation de phase) est plus petit. Ainsi, la ségrégation était considérée comme "parfaite" si seulement deux groupes distincts étaient formés à la fin du processus (les "phases" étaient complètement séparées).

Nous avons montré que même si les sujets recevaient tous, dans toutes les conditions testées, le même type d'information (un bip ou pas de bip), ils étaient inconsciemment sensibles à des variations subtiles de l'information traitée en amont par le système. Ainsi, le temps de ségrégation a une dépendance en forme de U par rapport à k, avec une valeur optimale à $k = 7 \sim 9$. Le nombre de groupes diminue, et la fraction de ségrégations parfaites augmente avec k, jusqu'à saturation à $k = 7 \sim 9$ également.

Nous avons ensuite utilisé le modèle pour prédire comment ces résultats seraient affectés par des changements dans la taille du groupe (groupes de 10, 22, 30 et 42 individus). Nous avons trouvé que la valeur optimale du temps de ségrégation ne changeait pas avec la taille du groupe, alors que le plateau (saturation) et la valeur optimale de k pour la qualité du clustering augmentaient tous deux avec la taille du groupe, de façon non linéaire. La fraction de ségrégations parfaites diminue quant à elle avec la taille de groupe. Le modèle, calibré pour la condition *majoritaire*, a pu reproduire les résultats obtenus pour les deux autres conditions, soulignant ainsi sa généralité.

L'analyse des environnements *exclusif* et *shifted* a montré de façon intéressante qu'avec la "complexité" de la tâche, le temps de ségrégation augmente, mais la qualité du clustering s'améliore. En particulier et de manière tout à fait inattendue, si l'on utilise les informations de trois voisins (condition *décalée*), ceux à considérer pour optimiser le processus de clustering sont les troisième, quatrième et cinquième plus proches. Ce phénomène non-trivial est bien reproduit par notre modèle.

Ce projet n'est pas encore entièrement terminé. En particulier, les mécanismes qui sous-tendent la dépendance de la qualité de la ségrégation par rapport à la taille de groupe et les phénomènes observés dans les environnements *exclusif* et *décalé* ne sont toujours pas bien compris. En outre, la force d'interaction entre deux piétons est encore à l'étude.

Cependant, nos résultats démontrent déjà la possibilité de concevoir des systèmes de traitement de l'information aidant la prise de décision humaine, et soulignent que de tels systèmes peuvent être utilisés pour *nudger* les phénomènes collectifs émergeant des interactions humaines dans des directions souhaitées. Ils seraient particulièrement utiles dans les événements de masse où un comportement collectif destructeur (panique de foule) est toujours susceptible de se produire. Bien sûr, ils ouvrent également la porte à la manipulation, de sorte qu'un contrôle rigoureux et une totale transparence seraient nécessaires avant qu'ils puissent être mis en oeuvre.

Résumé en Français

Dans cette thèse, nous nous sommes intéressés à l'impact de la quantité et de la qualité de l'information échangée entre individus d'un groupe sur leurs performances collectives dans deux types de tâches bien spécifiques. Dans une première série d'expériences, les sujets devaient estimer des quantités séquentiellement, et pouvaient réviser leurs estimations après avoir reçu comme information sociale l'estimation moyenne d'autres sujets. Nous contrôlions cette information sociale à l'aide de participants virtuels (dont nous contrôlions le nombre) donnant une information (dont nous contrôlions la valeur), à l'insu des sujets. Nous avons montré que lorsque les sujets ont peu de connaissance préalable sur une quantité à estimer, (les logarithmes de) leurs estimations suivent une distribution de Laplace. La médiane étant un bon estimateur du centre d'une distribution de Laplace, nous avons défini la performance collective comme la proximité de la médiane (du logarithme) des estimations à la vraie valeur. Nous avons trouvé qu'après influence sociale, et lorsque les agents virtuels fournissent une information correcte, la performance collective augmente avec la quantité d'information fournie (fraction d'agents virtuels). Nous avons aussi analysé la sensibilité à l'influence sociale des sujets, et trouvé que celle-ci augmente avec la distance entre l'estimation personnelle et l'information sociale. Ces analyses ont permis de définir 5 traits de comportement : garder son opinion, adopter celle des autres, faire un compromis, amplifier l'information sociale ou au contraire la contredire. Nos résultats montrent que les sujets qui adoptent l'opinion des autres sont ceux qui améliorent le mieux leur performance, car ils sont capables de bénéficier de l'information apportée par les agents virtuels. Nous avons ensuite utilisé ces analyses pour construire et calibrer un modèle d'estimation collective, qui reproduit quantitativement les résultats expérimentaux et prédit qu'une quantité limitée d'information incorrecte peut contrebalancer un biais cognitif des sujets consistant à sous-estimer les quantités, et ainsi améliorer la performance collective. D'autres expériences ont permis de valider cette prédiction.

Dans une seconde série d'expériences, des groupes de 22 piétons devaient se séparer en clusters de la même "couleur", sans indice visuel (les couleurs étaient inconnues), après une courte période de marche aléatoire. Pour les aider à accomplir leur tâche, nous avons utilisé un système de filtrage de l'information disponible (analogue à un dispositif sensoriel tel que la rétine), prenant en entrée l'ensemble des positions et couleurs des individus, et retournant un signal sonore aux sujets (émit par des tags attachés à leurs épaules) lorsque la majorité de leurs k plus proches voisins était de l'autre couleur que la leur. La règle consistait à s'arrêter de marcher lorsque le signal stoppait. Nous avons étudié l'impact de diverses valeurs de k sur le temps et la qualité de la ségrégation, définie comme le nombre de clusters à l'instant final, par analogie avec les phénomènes de séparation de phase (une ségrégation "parfaite" correspondant à la formation de deux clusters bien distincts). Nous avons trouvé que le temps de ségrégation est optimisé pour $k = 7 \sim 9$, et que la qualité de la ségrégation augmente avec k jusqu'à $k = 7 \sim 9$ également, valeur au-delà de laquelle elle sature. Notre dispositif nous a également permis d'enregistrer les positions des piétons durant les expériences, ce qui nous a permis de caractériser et modéliser les interactions des piétons avec le bord de l'arène et entre eux durant la marche aléatoire. À l'aide d'une procédure de minimisation d'erreur, nous avons reconstruit les formes fonctionnelles précises des interactions et construit un modèle fin de mouvement collectif de piétons.

Summary in English

In this thesis, we were interested in the impact of the quantity and quality of information exchanged between individuals in a group on their collective performance in two very specific types of tasks. In a first series of experiments, subjects had to estimate quantities sequentially, and could revise their estimates after receiving the average estimate of other subjects as social information. We controlled this social information through virtual participants (which number we controlled) giving information (which value we controlled), unknowingly to the subjects. We showed that when subjects have little prior knowledge about a quantity to estimate, (the logarithms of) their estimates follow a Laplace distribution. Since the median is a good estimator of the center of a Laplace distribution, we defined collective performance as the proximity of the median (log) estimate to the true value. We found that after social influence, and when the information provided by the virtual agents is correct, the collective performance increases with the amount of information provided (fraction of virtual agents). We also analysed subjects' sensitivity to social influence, and found that it increases with the distance between personal estimate and social information. These analyses made it possible to define five behavioral traits: to keep one's opinion, to adopt that of others, to compromise, to amplify social information or to contradict it. Our results showed that the subjects who adopt the opinion of others are the ones who best improve their performance because they are able to benefit from the information provided by the virtual agents. We then used these analyses to construct and calibrate a model of collective estimation, which quantitatively reproduced the experimental results and predicted that a limited amount of incorrect information can counterbalance a cognitive bias that makes subjects underestimate quantities, and thus improve collective performance. Further experiments have validated this prediction.

In a second series of experiments, groups of 22 pedestrians had to segregate into clusters of the same "color", without visual cue (the colors were unknown), after a short period of random walk. To help them accomplish their task, we used an information filtering system (analogous to a sensory device such as the retina), taking all the positions and colors of individuals in input, and returning an acoustic signal to the subjects (emitted by tags attached to their shoulders) when the majority of their k nearest neighbors was of a different color from theirs. The rule was to stop walking when the signal stopped. We studied the impact of various values of k on segregation time and quality, defined as the number of clusters at final time, by analogy with phase separation phenomena (a segregation was considered "perfect" when two distinct clusters were formed). We found that segregation time is optimized for $k = 7 \sim 9$, and that segregation quality increases with k up to $k = 7 \sim 9$ as well, value beyond which it saturates. Our device has also allowed us to record the positions of the pedestrians during the experiments, which allowed us to characterize and model the interactions of pedestrians with the border of the arena and between them during the random walk phase. Using an error minimization procedure, we were able to reconstruct the precise functional forms of the interactions and built a fine model of collective pedestrian motion.